Astronomy Unit
School of Physics and Astronomy
Queen Mary University of London

Formation and Dynamical Evolution of Planetary Systems

Sanson Tsun Sum Poon

Supervised by

Richard P. Nelson

Submitted in partial fulfilment of the requirements of the Degree of
Doctor of Philosophy

Declaration

I, Sanson T. S. Poon, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated. Previously published material is also acknowledged below.

I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material.

I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis.

I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university. The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without the prior written consent of the author.

Details of collaboration and publications:

Sanson T. S. Poon, Richard P. Nelson, Seth A. Jacobson and Alessandro Morbidelli (2020).
Formation of compact systems of super-Earths via dynamical instabilities and giant impacts. MNRAS 491, 5595-5620

Sanson T. S. Poon and Richard P. Nelson (2020).
On the origin of the eccentricity dichotomy displayed by compact super-Earths: dynamical heating by cold giants. MNRAS 498, 5166-5182

Sanson T. S. Poon, Richard P. Nelson and Gavin A. L. Coleman (submitted).
In situ formation of hot Jupiters with companion super-Earths.

Signature: Sanson T. S. Poon
Date: February 19, 2021

In memory of

Chan Lin

Abstract

Our understanding of extrasolar planetary systems has undergone a revolution in recent years due to major advances in the observations of exoplanets. Planetary system architectures are found to be very diverse, consisting of closely orbiting compact multiple systems of super-Earths and terrestrial planets, widely spaced systems of giant planets, and bodies on temperate orbits at intermediate distances from their host stars. A major unsolved question is how did these planetary systems form and evolve to create the diversity that we now observe?

The focus of this thesis is to address this question using N-body simulations of planetary systems, combined with models for realistic planet-planet collisions and the nascent protoplanetary disc with prescriptions for disc-planet interactions. The aim is to adopt initial conditions similar to those thought to exist around young stars that are in the process of forming, and to construct a synthetic population of exoplanet systems for comparison with the observations, in order to test our current theories of planetary system formation. Different planetary system formation and evolution scenarios are examined in this thesis, including *in situ* formation and self-scattering between growing protoplanets, dynamical heating of inner planetary systems by cold gas giants, and the *in situ* formation of systems containing hot Jupiters and interior super-Earths.

In the *in situ* self-scattering scenario, simulations show that the final planetary systems look broadly similar to the compact multiplanetary systems observed by Kepler. But the observed distributions of planetary multiplicities or eccentricities are not reproduced, because scattering does not excite the systems sufficiently. Post-processing the collision outcomes suggests that the planets would not significantly change the ice fractions of initially ice-rich protoplanets, but significant stripping of gaseous envelopes appears likely. Hence, it may be difficult to reconcile the observation that many low-mass Kepler planets have H/He envelopes with an *in situ* formation scenario that involves giant impacts after the dispersal of the gas

disc.

In the scenario of dynamical heating of the inner planetary system by cold giants, synthetic transit observation of the final systems reveals dichotomies in both the eccentricity and multiplicity distributions that are close to being in agreement with the Kepler data. This suggested that understanding the observed orbital and physical properties of the compact systems of super-Earths discovered by Kepler may require holistic modelling that couples the dynamics of both inner and outer systems of planets during and after the epoch of formation.

The *in situ* formation scenario demonstrates the plausibility of forming systems of a hot Jupiter with companion super-Earths, such as WASP-47, Kepler-730, and TOI-1130. Simulations show the evolution for such systems consistently follows four distinct phases: early giant impacts; runaway gas accretion onto the seed protoplanet; disc damping-dominated evolution of the embryos orbiting exterior to the giant; a late chaotic phase after the dispersal of the gas disc. Synthetic transit observations of the simulations provide a similar occurrence rate for systems containing a hot Jupiter and an inner super-Earth to the actual transit surveys. But simulated hot Jupiters are rarely detected as single transiting planets, in disagreement with observations. *In situ* formation is a viable pathway for forming systems with these unusual architectures, although not for the majority of hot Jupiters.

Acknowledgements

雲蒼蒼 茫茫長夜裡感激 有空港
雲層上滿星光 星星不知道我的怯慌
人間卻有殷切的眼光
每夜 往上看

— 容祖兒, 黃偉文 ⟨空港⟩

This thesis would not have been possible without the guidance and enlightenment of my supervisor, Richard Nelson. I am sincerely grateful for his time and patience throughout my time in the Astronomy Unit. I would also like to thank my second supervisor, Sijme-Jan Paardekooper, for all his support and constructive discussion. A special thanks to Jim Emerson, who provided me many helpful suggestions to choose my marvellous supervisors and research area.

I have had a fantastic and enjoyable time during my Ph.D. study in the Astronomy Unit. I would like to thank everyone that I came across in the last four years, especially but not only to (in no particular order):

John Ronayne and Shailee Imrith, who always are very good company in the office during all the overtimes, evenings and weekends; Maritza Soto, for all her interesting conversation, and our unforgettable road trip in Iceland; Pedro Carrilho and Charalambos Pittordis, for all the amazing barbecue parties; Viraj Sanghai, for all your laughs; Colin McNally, who has knowledge in almost everything both inside and outside his research area, including if Jesus wear socks in the sandals; Francesco Lovascio, for putting our photo as his online profile picture, and the enjoyable conference trip in Japan and Iceland; Jack Skinner, for our amusing discussions during the coffee time; Sandy Zeng, the only person that I can chat to with my native language; Domenico Trotta, for the delicious spaghetti carbonara; Jorge Venegas, for sharing my feeling when the football team that we support lost almost every week; Clark Baker, who always bring fun to everyone in the office; Louis Coates, for the free ice-skating entrance; and Christopher Gallagher, for his great sense of humour.

Also, all the lecturers who gave me lectures in my Astrophysics course, including

Carl Murray, Guillem Anglada-Escude, James Cho, David Burgess, Will Sutherland, Sergei Vorontsov, and Alexander Polnarev; every member in the planetary group for many constructive discussions, including Craig Agnor, Nick Cooper, Richard Donnison, Edward Gillen, Thomas Haworth, Izaskun Jimenez-Serra, Gavin Coleman, David Quenard, Ali Barlas, Kevin Chan, Paul Hallam, Matthew Mutter, John Strachan, and George Turpin, from whom I learned a lot; the cosmology group and plasma group for all the lunch-time and coffee-time chats, including Tessa Baker, Phil Bull, Bernard Carr, Christopher Chen, Chris Clarkson, Timothy Clifton, David Mulryne, Alkistis Pourtsidou, Ian Roxburgh, David Tsiklauri, Julian Adamek, Samir Choudhuri, Luca Franci, Callum Boocock, Jesse Coburn, Eline De Weerd, Jessie Durk, Alice Giroul, Fraser Kennedy, Md Goribullha Shah, Paula Soares, and specially Karim Malik, our department head and the second tallest member of the Astronomy Unit; and my friendly and energetic colleagues at Royal Observatory Greenwich: Patricia Skelton, Gregory Brown, Dhara Patel, Angelos Tsiaras, Tania de Sales Marques, Anna Ross, Elizabeth Avery, Emily Drabek-Maunder, Edward Bloomer, Mikaela Webb, and Bryony Lanigan.

A big and special thanks to Rebeca Martinez-Carrillo, who always supports me and makes my research time extra enjoyable.

Last but not least, I am indebted to my parents and close family: Tim, San, and Amy, for all their patience and continuous support.

Contents

1. Introduction to exoplanets

一閃一閃亮晶晶 好像你的身體
藏在衆多孤星之中還是找得到你

— 孫燕姿, Hush 〈克卜勒〉

1.1. A brief history of exoplanets

Contemplating the existence of an unknown world, rather than the one we are living on, has always fascinated human beings. The conjecture that our Earth and the solar system are not the only ones in the universe has always emerged in history. To date, planets beyond our solar system can be detected. They are named exoplanets. No matter how we define exoplanets, either as worlds or planets, curiosity about them is a natural part of the human instinct.

1.1.1. From the ancient periods to the first detection

As early as the classical period of ancient Greek history (c. 600 B.C. - c. 300 B.C.), philosophers had already thought that there are some other worlds existing other than our own. For example, Anaximander's (c. 610 B.C. - c. 546 B.C.) assumption of the worlds to be infinite in numbers was mentioned by Simplicius of Cilicia (c. 490 - c. 560) in his work *On Aristotle's Physics*, and Epicurus (341 B.C. - 270 B.C.) in his *Letter to Herodotus* said:

"*There are infinite worlds both like and unlike this world of ours.*".

Their comments show that the belief in the existence of another world had emerged a long time ago and is not a new idea in the recent decade.

Hippolytus of Rome (170 - 235) remarking on Democritus (c. 460 B.C. - c. 370 B.C.) said:

"*There are innumerable worlds, which differ in size. In some worlds there is no sun and moon, in others they are larger than in our world,*

> *and in others more numerous. The intervals between the worlds are unequal, in some parts there are more worlds, in others fewer, some are increasing, some at their height, some decreasing, in some parts they are arising, in others failing. They are destroyed by collision one with another. There are some worlds devoid of living creatures or plants or any moisture.*".

This provided a more detailed description of the ancient idea. If we try to understand this statement from the modern astrophysics point of view, it included some important aspects which can be connected to our understanding of planetary system structure and evolution. It mentioned that worlds can be different in size, and we know that the type of planets can be very diverse (e.g. terrestrial planets, ice giants, and gas giants). For the number of suns and moons, some planets can have more than one satellite (e.g. Saturn) and some can have none (e.g. Venus), and some planets are orbiting not only a single star (e.g. in a binary star system). The number of worlds can be different in the sense of different multiplicities of different exoplanetary systems once observed. The change of number of worlds can be described as the evolution process of a planetary system, where the nebular hypothesis (section 1.4) suggested that many planets can be formed in a system and the planets grow from the settling of dust grains to form planetesimals, planetary embryos, protoplanets, and finally to planets. At some stage of the evolution process, the high number of protoplanets/planets might experience collisions and giant impacts with different bodies, which would then result in decreasing the multiplicity of the system. Also, some exoplanets are discovered at an 'optimum' distance away from their host star where the planet surface temperature is in the range that can support liquid water (habitable zone). Earth-like life may then develop. Although the ancients thought and the belief in 'another world' is very likely to be different from the modern physical model, it still provided evidence that people are curious about what is outside the Earth.

Those ideas were not commonly accepted, and an influential philosopher in that period of time, Aristotle (384 B.C. - 322 B.C.), did not agree with the idea. Aristotle in his work, *On the Heavens*, asserted that

> "*There cannot be more worlds than one.*".

In the 2nd century, the notion on existence of 'multiple world systems' emerged in Asia. An Indian Buddhist philosopher, Nagarjuna (c. 150 - c. 250), mentioned that there are 10 billion '*Sumeru*' (centre of a world) and 10 billion suns and moons

together to form a bigger world system (can be thought of as a galaxy) in his book *The Treatise on the Great Perfection of Wisdom*. Although he did not provide any proof of his work in a scientific way, this idea was commonly shared in the area of the Buddhist cultural sphere in that period of time.

Later in the period of Renaissance (the 14th - 17th centuries) in Europe, the conjecture of infinite numbers of worlds transpired again. An Italian philosopher, Giordaro Bruno (1548 - 1600), said:

"*There are an infinity of worlds of the same kind as our own.*".

A major landmark was made by Galileo Galilei (1564 - 1642) as he first observed other planets in the solar system with a telescope. This provided a strong support to Nicolaus Copernicus's (1473 - 1543) idea of heliocentrism and opened a door for searching for other 'worlds'. Due to the limitation in technology at that time, the first actual search of exoplanets was not made until the 19th century.

1.1.2. From the first detection to present

In the middle of 19th century, the first exoplanet detection was made. Jacob (1855) gave some signs of a planetary object being present around the binary star 70 Ophiuchi. Four decades later in 1896, another observation of 70 Ophiuchi was made by See (1896) where he tried to predict the existence of a dark object with a period of 36 years due to the unexpected orbital motion of the systems. However the system predicted by See (1896) is thought to be unstable (Sherrill, 1999).

In 1969, a slight unsteady movement of Barnard's star (a very low mass and faint star near β Ophiuchi) is reported in Van de Kamp (1969) and claimed a discovery of an orbiting planet. Although the above claim was based on the wrong information (Boss, 2009), they took the first step to extrasolar planet exploration.

With the technological improvement in recent decades, the detection of exoplanets can be confirmed. The first detection of an extrasolar planet was made in 1989 by the radial velocity method (HD 114762 b, Latham et al., 1989), which was confirmed by Cochran et al. (1991)[1]. Subsequently in 1992, two super-Earths, PSR B1257+12c and PSR B1257+12d, were discovered by Wolszczan and Frail (1992). They are orbiting a pulsar (pulsating radio star) and detected by the pulsar timing

[1]A recent Gaia data analyses by Kiefer et al. (2019) and Kiefer (2019) suggested that the mass of HD 114762 b is ~ 100 M_J, and is far beyond the planet domain. Although an object in this range of mass is usually classified as a brown dwarf, HD 114762 b is considered as an exoplanet according to the NASA Exoplanet Archive.

method. Also, the PSR B1257+12 system is the first multiple extrasolar planet system discovered.

The first confirmed extrasolar planet orbiting around a main-sequence star, 51 Pegasi (with stellar classification of G5V), is 51 Pegasi b (Mayor and Queloz, 1995). The mass of 51 Pegasi b has a lower limit about half of Jupiter's mass, and the semimajor axis of its orbit is about 0.05 au. According to its value of mass and semi-major axis, 51 Pegasi b is classified as a hot Jupiter. The first confirmed super-Earth orbiting around a main sequence star is Gliese 876 d which is found by Rivera et al. (2005).

The development in the field of searching for exoplanets did not slow down because of the early success. For example, the usage and improvement of charge-coupled device (CCD) cameras, image processing computer software and spectroscopy with high resolution, have led to some new detection methods such as transit and direct imaging. And also, with the improved measurement and analysis procedure, more planet candidates can be confirmed as planets, such as γ Cephei Ab which was detected in 1988 (Campbell et al., 1988) and confirmed 15 years later in 2003 (Hatzes et al., 2003).

Furthermore, the culture of searching for exoplanets has changed from individual-based to mission- and project-based in the past decade. Convection Rotation and planetary Transits (CoRoT, Auvergne et al., 2009) led by National Centre for Space Studies (CNES) is the first extrasolar planet space mission which was launched in 2006. CoRoT detected planets by the transit method and obtained more than 160,000 light curves in about 7 years. After the launch of CoRoT, more and more extrasolar planet missions (both ground based project and space mission) were introduced, such as Kepler (Borucki et al., 2010), Wide Angle Search for Planets (WASP, Pollacco et al., 2006), and High Accuracy Radial Velocity Planet Searcher (HARPS, Mayor et al., 2003).

1.2. Detection methods

The detection of extrasolar planets is not as easy as the detection of stars or galaxies. Even in the case of a planet with a high albedo, the light from the host star that is reflected by the planet is still extremely faint compared to the light that is received directly from the host star. Not only is the detection of such a tiny amount of light difficult, but also the relatively bright light from the host star has almost covered all the light source information of the planet. This situation can be analogous to

observing a firefly flying close to a lighthouse. The light from the firefly are still able to reach the observer, but it is extremely difficult to classify the actual source within the lightning bug (planet) and the lighthouse (star).

With the above reason and the technology standard nowadays, most of the extra-solar planet detection methods are indirect observation. Some orthodox methods, such as radial velocity and transit, will be discussed in this section.

1.2.1. Radial velocity method

The radial velocity (RV) method is also known as Doppler spectroscopy. The Doppler effect is the change in frequency of a wave when the source moves relative to its observer along the line of sight. Compared to the emitted frequency from the source, the observed frequency is higher during the approach and lower during the recession. The idea of using the radial velocity method to search for exoplanets goes back to the 1950's (e.g. Struve, 1952). As of today, around 20% of exoplanets were discovered by this method. The first confirmed planet around a main-sequence star (Mayor and Queloz, 1995) was discovered by the radial velocity method. HARPS on La Silla Observatory and the High Resolution Echelle Spectrometer on Keck Observatory (HIRES, Vogt et al., 1994) are two of the highest precision ground-based instruments and most successful projects that use Doppler spectroscopy to find exoplanets. More than 400 exoplanets were found by these instruments, including one of the smallest RV exoplanets, YZ Cet b ($< 1\ M_\oplus$), orbiting around a nearby M-dwarf YZ Cet (Astudillo-Defru et al., 2017).

When a planet is orbiting a star, the star would feel the gravitational force from the planet. The gravitational force from the planet would induce a small periodic orbital motion to the star around their barycentre (figure 1.1) due to the conservation of momentum. Detecting such a periodic wobble of the star suggested that there might be a planet present in that system. This wobbling motion of a star can be detected and analysed by spectroscopy. As shown in figure 1.2, the starlight is blue-shifted (shorter wavelength) when the star moves towards and red-shifted (longer wavelength) when the star moves away the observer.

The Doppler shift of the star caused by the planetary orbital motion can be expressed by the Doppler formula,

$$\frac{\lambda_{\rm obs} - \lambda_{\rm emit}}{\lambda_{\rm emit}} = \frac{v_{\rm R} + v_\star}{c}, \tag{1.1}$$

where $\lambda_{\rm obs}$ is the wavelength detected by the observer, $\lambda_{\rm emit}$ is the emitting wave-

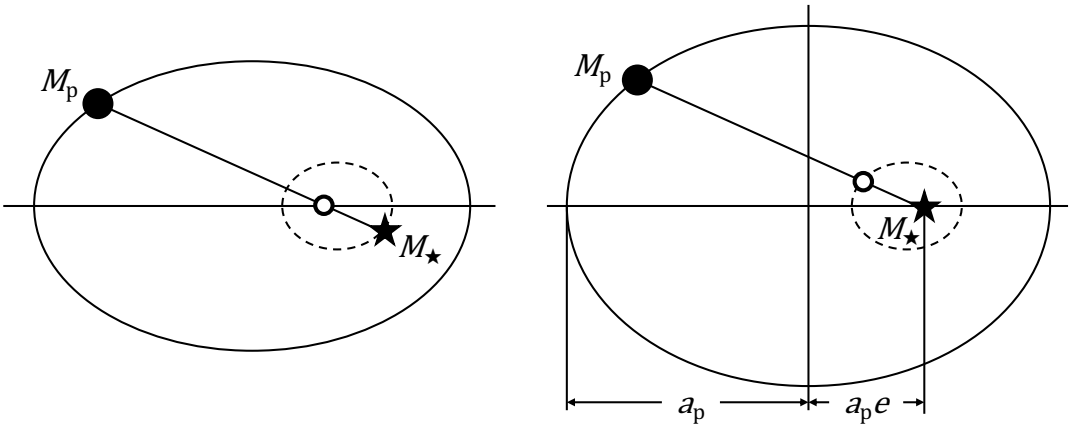

Figure 1.1.: Two-body planetary system with their orbits around the barycentre (∘) of the system. The solid orbits represent the orbits of mass M_p around the barycentre (left panel) and the central mass $M_\star$ (right panel). The dashed orbits are the orbit of the mass $M_\star$ around the barycentre (left panel) and the orbit of the barycentre in the reference frame of the centre mass (right panel). The orbit of the planet has a semi-major axis a_p and an eccentricity e about the star. Figure not to scale.

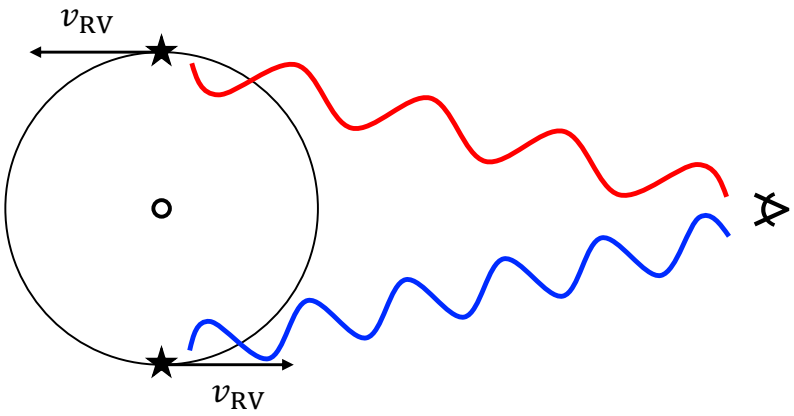

Figure 1.2.: Top view of a radial velocity model. The wavelength of the light from the star gets shorter when the motion of the star is moving towards the observer and gets longer when the motion is moving away from the observer. The symbol ∘ is the centre of the stellar orbit and the arrow shows the direction of the star's orbit. Figure not to scale.

length of the source star, $v_{\rm R}$ is the relative radial speed caused by the planet orbital motion between the star and the observer, $v_\star$ is the relative speed between the frame of the star and the observer, and c is the speed of light.

With the known value of $v_{\rm R}$, we can then derive the RV signal (velocity semi-amplitude), $v_{\rm RV}$ by,

$$v_{\rm RV} = \frac{\max\left(v_{\rm R}\right) - \min\left(v_{\rm R}\right)}{2}. \tag{1.2}$$

Hence, the conservation of momentum between the star and the planet yields (Seager, 2010; Lovis and Fischer, 2010),

$$v_{\rm RV} = \left(\frac{2\pi G}{P}\right)^{1/3} \frac{M_{\rm p} \sin i_{\rm p}}{\left(M_{\rm p} + M_\star\right)^{2/3}} \frac{1}{\left(1 - e^2\right)^{1/2}}, \tag{1.3}$$

where G is the gravitational constant, P is the orbital period, $M_{\rm p}$ is the mass of the planet, $M_\star$ is the mass of the star, e is the eccentricity of the orbit, and $i_{\rm p}$ is the inclination angle between the sky plane of the observer and the orbital plane of the planet. This formula can also be expressed in a more practical set of units,

$$v_{\rm RV} \approx \frac{28.4}{\left(1 - e^2\right)^{1/2}} \left(\frac{P}{1\ \rm yr}\right)^{-1/3} \left(\frac{M_{\rm p} \sin i_{\rm p}}{\rm M_J}\right) \left(\frac{M_{\rm p} + M_\star}{\rm M_\odot}\right)^{-2/3} \rm ms^{-1}, \tag{1.4}$$

where the unit $\rm M_J$ is the Jupiter mass and $\rm M_\odot$ is the solar mass.

Figure 1.3 shows two examples for the radial velocity detections of planets. The radial velocity curve on the top panel is the discovery detection of the first confirmed planet around a main-sequence star, 51 Pegasi b (Mayor and Queloz, 1995), which is a $M_{\rm p} \sin i_{\rm p} \sim 0.47$ $\rm M_J$ giant with an orbital period of 4.23 days around a 1.03 $\rm M_\odot$ star. This hot Jupiter provides a strong RV signal of $v_{\rm RV} \sim 60$ $\rm ms^{-1}$. The curve on the bottom panel is the discovery detection of Proxima Centauri b, which is orbiting the nearest star to our Solar system and is only 4.2 light-year away (Anglada-Escudé et al., 2016). Proxima Centauri is a red-dwarf with $M_\star = 0.12$ $\rm M_\odot$, and the mass of its planet has the $M_{\rm p} \sin i_{\rm p}$ value ~ 0.004 $\rm M_J$ (1.27 $\rm M_\oplus$). This yields a small RV signal of $v_{\rm RV} \sim 1.38$ $\rm ms^{-1}$.

One drawback of the radial velocity method is that the detection of low mass planets around a Sun-like star is difficult and requires high RV detection accuracy. It is because the values of $v_{\rm RV}$ for such systems are too small for detection limits nowadays (generally $\sim$ 1 to 3 $\rm ms^{-1}$). More recently, ESPRESSO (instrumental precision ~ 0.1 $\rm ms^{-1}$, Pepe et al., 2020) demonstrated a RV precision of ~ 0.25 $\rm ms^{-1}$ (limited by stellar noise). Table 1.1 lists the RV signal of the Sun induced by different

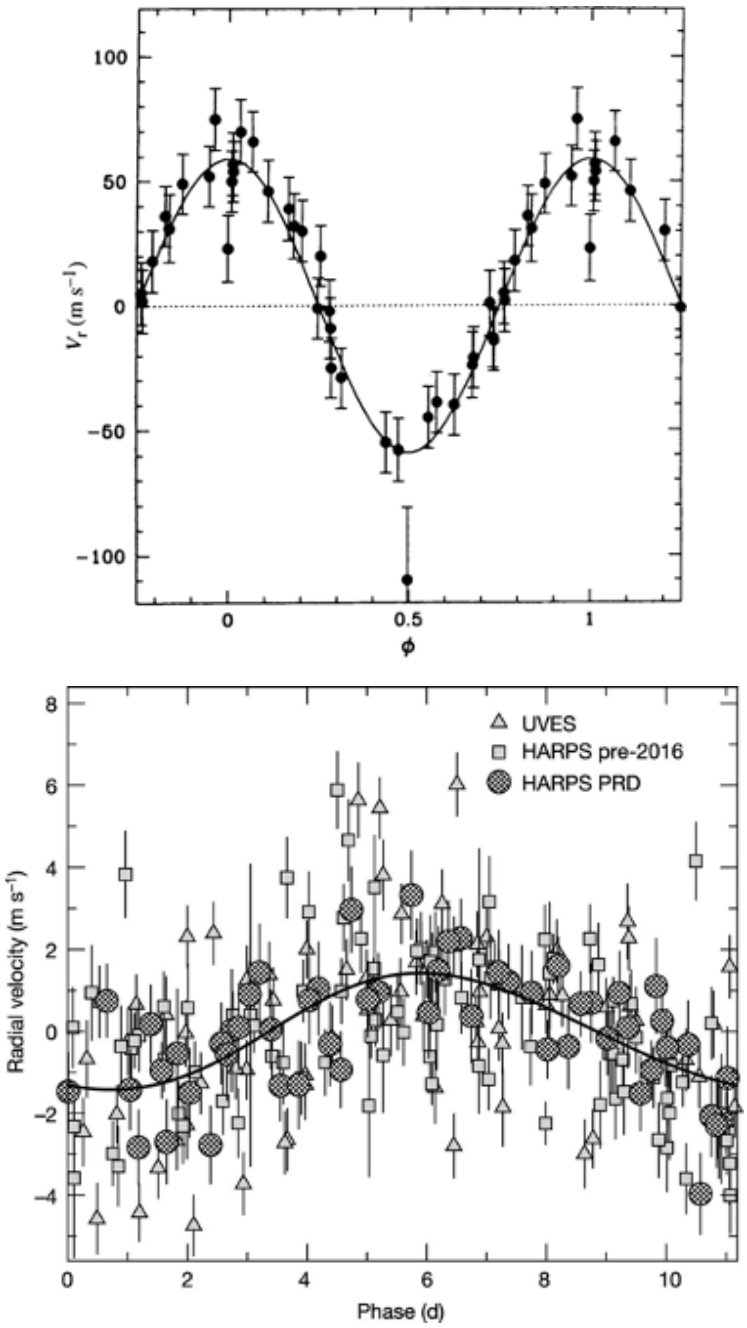

Figure 1.3.: The discovery RV measurements of (top) 51 Pegasi b (Mayor and Queloz, 1995), and (bottom) Proxima Centauri b (Anglada-Escudé et al., 2016) with their model fits plotted in solid lines. The x-axes repersent the phase, where (top) ϕ is the time normlized to one epoch based on their model fit, and (bottom) is the real time in the unit of days. The data points are obtained by (top) the ELODIE of the Haute-Provence Observatory (Baranne et al., 1996), and (bottom) the UVES (Dekker et al., 2000) and HARPS both at the European Southern Observator (ESO).

Table 1.1.: RV signal, v_{RV} of different planets with an edge on circular orbit around a Solar type star ($i_p = 90°$; $e = 0$; $M_\star = 1M_\odot$). Values of v_{RV} are obtained by equations 1.3 and 1.4.

Planet	Semi-major axis [au]	v_{RV} [$m\,s^{-1}$]
Jupiter	5.2	12.5
Saturn	9.6	2.75
Uranus	19.2	0.30
Neptune	30.1	0.28
Earth	1.0	0.089
Venus	0.7	0.087
Mercury	0.4	0.008
Mars	1.5	0.008
Jupiter	0.1	89.8
Neptune	0.1	4.8
Earth	0.1	0.28
Jupiter	1.0	28.4
Neptune	1.0	1.5

solar system planets. In the solar system, only Jupiter lies above the RV detection limit ($v_{RV} = 12.5$ ms^{-1}) while Saturn is close to the limit. A Neptune size planet would also be detectable if it is orbiting much closer to the host star (at 0.1 au, $v_{RV} = 4.8$ ms^{-1}). To detect planets with smaller values of v_{RV}, higher precision RV observations are needed, such as the future Thirty Metre Telescope (TMT, Schöck et al., 2009) and its installed ExAO systems which might be able to observe Neptune-like planets orbiting beyond 1 au.

1.2.2. Transit method

If a planet transits between its host star and the observer, the planet would block some of the light from its star. Therefore, if we detect a periodic drop of the stellar brightness, it suggests that there might be a planet orbiting around the star. Transit events can be detected by monitoring the changes in the brightness of the star. During a transit event, the relative brightness of the star would drop and this could be recorded in a light curve (figure 1.4). Figure 1.5 shows two examples of a transit light curve obtained from the Kepler observations together with their transit model fits.

The period of the planet can be obtained by measuring the time in between each

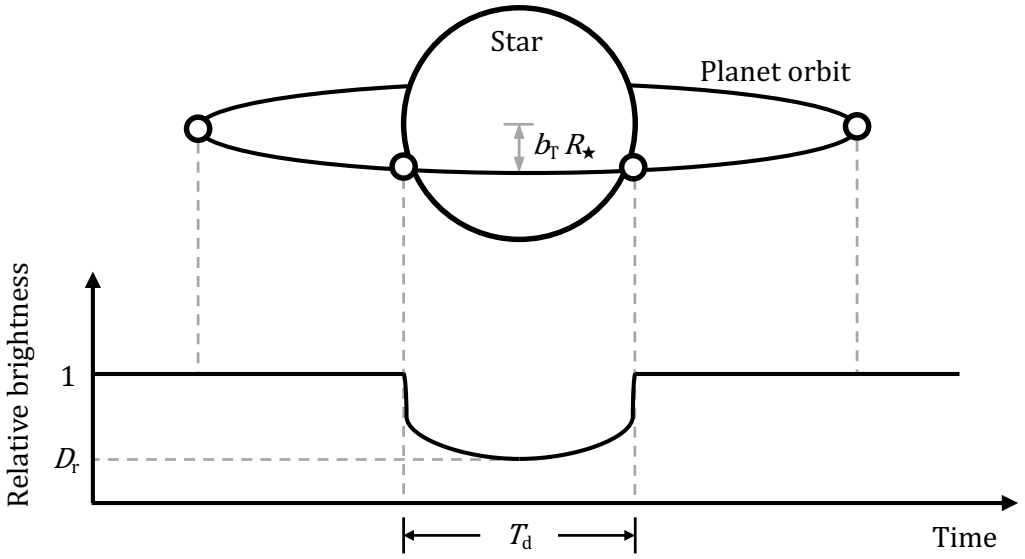

Figure 1.4.: Diagram illustrating a transit event. At the bottom is the light curve of the relative brightness of the star. The transit impact parameter is b_T (see also figure 1.6), the radius of the star is $R_\star$, the relative transit depth is D_T, and the transit duration is T_d.

transit event, and the semi-major axis, a, can be calculated by the Kepler's 3rd Law,

$$a^3 = \frac{GM_\star}{4\pi^2} P^2, \tag{1.5}$$

when the stellar mass is also known. The transit light curve can also provide information of the planet radius, R_p, by the relation between the flux, F, and the area A, in the form of $F \propto 1/A$. Knowing the relative transit depth, D_T, the planet radius can be calculated by,

$$\frac{R_p}{R_\star} = \sqrt{\frac{\Delta F}{F}} = \sqrt{D_T}, \tag{1.6}$$

where $R_\star$ is the stellar radius and ΔF is the change of flux during the planet transit. If the planet masses can be found by follow-up RV observations, the density of the planets can also be determined. When the planet is transiting, the starlight would pass through the atmosphere of the planet and lead to some changes in the observed spectrum. The changes in the spectrum might be able to provide information on the planet's atmosphere, such as a recent discovery of water vapour in the atmosphere of K2-18b (Tsiaras et al., 2019).

The orbital eccentricity of a transit planet can also be estimated by the transit method. Figure 1.6 demonstrates the geometry required to obtain the eccentricity estimation. Knowing the values of P and a, we can calculate the expected transit

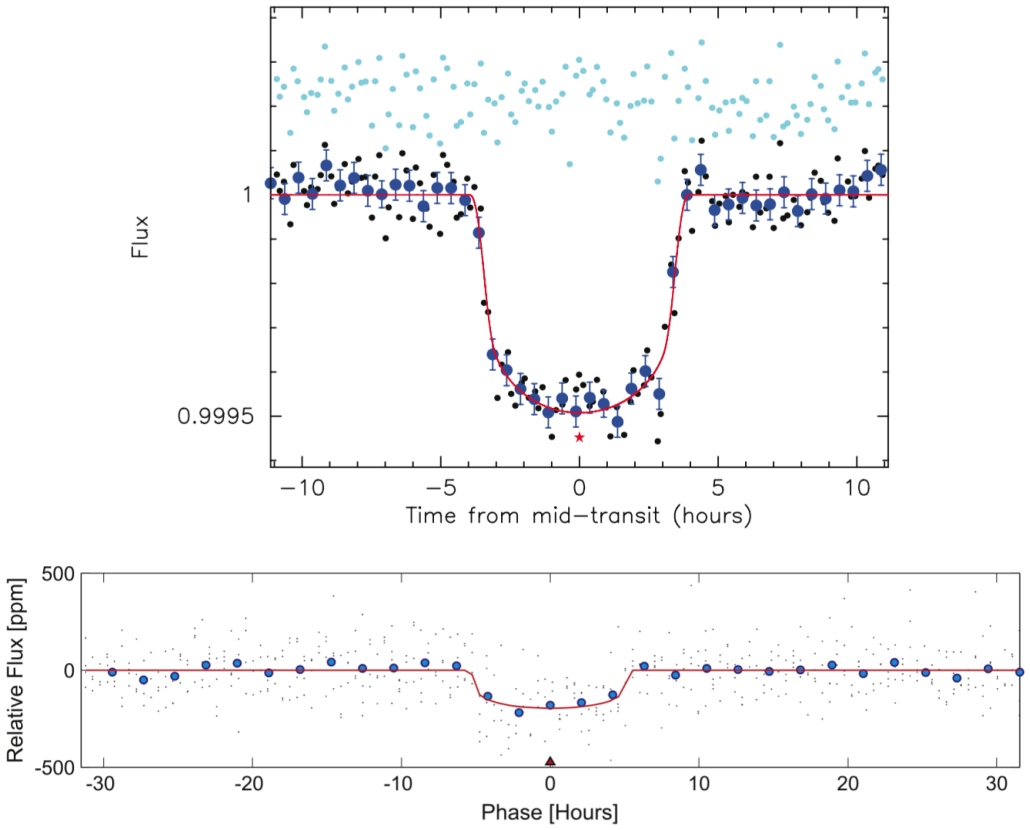

Figure 1.5.: Two examples of transit light curves. These are the discovery light curves of Kepler-22b (top) and Kepler-452b (bottom) with their model fits plotted in red line (Borucki et al., 2012; Jenkins et al., 2015). The red star (top) and triangle (bottom) symbols are their mid-transit times based on the model fit.

duration, $T_{\mathrm{d},0}$, while considering the planet in an imagined circular orbit by,

$$T_{\mathrm{d},0} = \frac{P}{2\pi}\varphi = \frac{P}{2\pi}\left[2\sin^{-1}\left(\frac{l}{a}\right)\right], \tag{1.7}$$

where φ is the transit duration angle defined as $\sin\varphi/2 = l/a$, l is the half distance the planet travelled across the surface of the star, $l = \sqrt{(R_\star + R_\mathrm{p})^2 - (b_\mathrm{T} R_\star)^2}$, b_T is the transit impact parameter which is defined as $b_\mathrm{T} = (a\cos i_\mathrm{p})/R_\star$, and i_p is the angle between the sky plane of the observer and the orbital plane of the planet. Together with the known transit duration, T_d, obtained directly from the observational light curve, the eccentricity of the orbit can be related to the argument of periastron, ω, by

$$\frac{T_\mathrm{d}}{T_{\mathrm{d},0}} = \frac{\sqrt{(1-b_\mathrm{T}^2)\,(1-e^2)}}{1+e\sin\omega}. \tag{1.8}$$

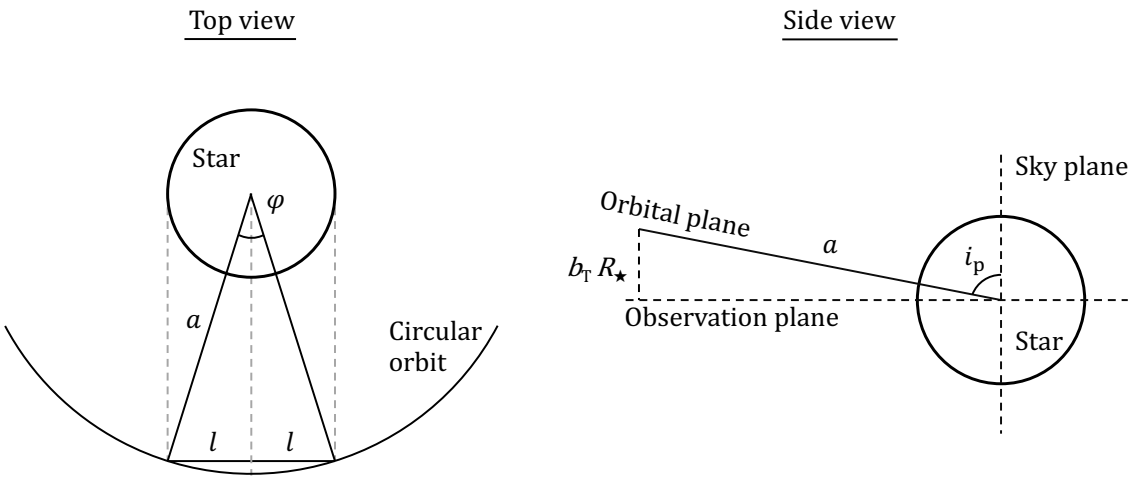

Figure 1.6.: Geometric demonstration of the transit model, defining the transit duration angle, φ, half distance the planet travelled across the surface of the star, l, and the transit impact parameter, $b_{\rm T}$.

The chances of detecting exoplanets by the transit method are generally very small, because not all planets have a transiting orbit from our line of sight. The probability of a planet having a transiting orbit, $P\,(\text{transit})$, can be approximated as,

$$P\,(\text{transit}) \approx \frac{R_\star}{a_{\rm p}}. \tag{1.9}$$

For example, when we are observing our solar system far away from a random position, the probability of seeing that the Earth transit the Sun would only be ~ 0.47 %. Moreover, even though the mutual inclination angle of the solar system is small ($< 4°$, except Mercury is $\sim 7°$, compared to the Earth's orbital plane), there is no single viewing angle that can allow all 8 planets to have a transiting orbit. From any viewing angle, the Solar system can only have a maximum number of 3 transiting orbits at a time. For any transit survey on their own, it is difficult to confirm the intrinsic multiplicities of those systems.

Although it's very likely that many planets are not transiting planets from our viewing position, the transit method has still discovered more than 75 % of confirmed planets to date. This is due to the fact that the transit survey can be targeting a large area of sky to achieve a relatively high efficiency for searching exoplanets, such as the CoRoT (Auvergne et al., 2009) and Kepler mission (Borucki et al., 2004). On the other hand, a significant weakness of transit surveys is the high degree of false positives (e.g. Santerne et al., 2012). Around half of the Kepler objects of interest (KOI) are considered as false positives. Many follow-up observations, such as RV, are required to validate those transit candidates from the transit surveys.

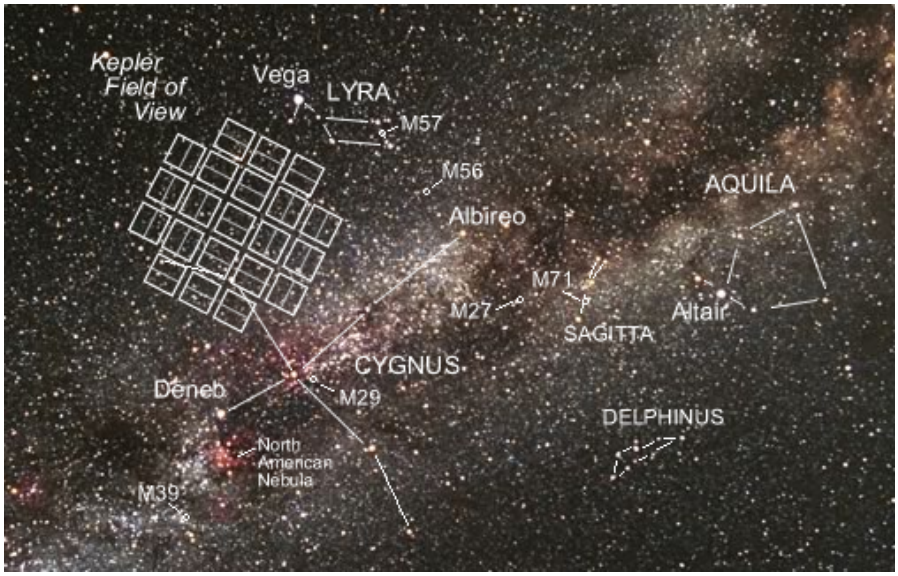

Figure 1.7.: Kepler field of view superimposed on the night sky (credit to C. Roberts; taken on 18-12-2009).

Kepler mission

The Kepler mission is designed to investigate the structure and variety of extrasolar planetary systems, and also to discover and validate Earth-size extrasolar planets in or near the habitable zone (Borucki et al., 2004). From the beginning of the mission, more than 2,300 planets have been confirmed as Kepler planets (Borucki et al., 2010, 2011; Batalha et al., 2013; Burke et al., 2014; Rowe et al., 2015; Mullally et al., 2015; Coughlin et al., 2016; Thompson et al., 2018). This makes the Kepler mission one of the most successful projects for finding exoplanets.

The Kepler Space Telescope is a space observatory that was launched by National Aeronautics and Space Administration (NASA) on the 7th March 2009. The main telescope of Kepler is a Schmidt type telescope with a diameter of 0.95 m and a collecting area of 0.708 m^2. Its orbit is a heliocentric orbit following the Earth with a semi-major axis near 1 au, so its field of view would not be covered by the Earth and the observations can be continuous, and the spectrophotometer will not be affected by the diffuse light from Earth. The Kepler field of view was pointing to the Cygnus constellation (figure 1.7), which is decent for the survey. It is because the view of Kepler to Cygnus would not be blocked by the Kuiper belt and the asteroid belt, and it is a dense region in the Galaxy (along the Orion spiral arm). Further technical overview of Kepler is set out in Borucki et al. (2004).

As mentioned earlier in this section, the confirmation of a Kepler planet is not only decided by the detection of a transit event. The validation process starts from

the Kepler Input Catalog (KIC, Brown et al., 2011), which contains the physical and photometric data of the stellar source in Kepler's field of view. If a periodic drop of brightness is detected from the source, that star is likely to host one or more transiting planets and to be considered as a member of the KOI. Once the star is in the list of KOI, Kepler would record more details on each transit event, such as the transit depth and transit duration of the signal. Follow-up observations are required to carry on the validation process (Gautier et al., 2010). With the follow-up observation, some transit events would become false positives (Morton et al., 2016) and some were confirmed as Kepler planets. In between the stage of becoming a confirmed planet or false positive, the KOI objects are classified as Kepler candidates, where they may or may not be a planet. The discovery statistic obtained from the Kepler mission is discussed further in section 1.3.2.

1.2.3. Other methods

Astrometry

Astrometry is the most common method that has been used to search for extrasolar planets from the 19th to the middle of 20th century. As mentioned earlier in section 1.1.2, Jacob (1855) and See (1896) did some observations of the 70 Ophiuchi system by astrometry. So far, only 1 confirmed planet was discovered using this detection method.

Similar to the RV method, astrometry also involved the observation on the orbital motion of the star due to its conservation of momentum with the planet (figure 1.1). While the RV method is to observe the Doppler shift of the stellar spectrum (section 1.2.1), astrometry observes the physical wobbling of the target star. Unfortunately, the changes of the stellar position are usually small, which is due to the planet-to-star mass ratio is usually small, and the gravitational force acted on the star by the planet is relatively tiny.

DENIS-P J082303.1-491201 b (Sahlmann et al., 2013) is the only exoplanet discovered by the astrometry method so far. This planet is also one of the most massive exoplanets found, and its mass is up to 28.5 M_J[2]. On the other hand, the mass of its host star (DENIS-P J0823031-491201) is only 0.075 $M_\odot$. Hence, the planet-to-star mass ratio of this system is big ($\sim$ 0.36, while Jupiter-to-Sun mass ratio is $\sim 9.5 \times 10^{-4}$), and the detection of the planet by astrometry is possible.

High precision observations are required for astrometric planet detections. The

[2]Although an object in this range of mass is usually classified as a brown dwarf, DENIS-P J082303.1-491201 b is considered as an exoplanet according to the NASA Exoplanet Archive.

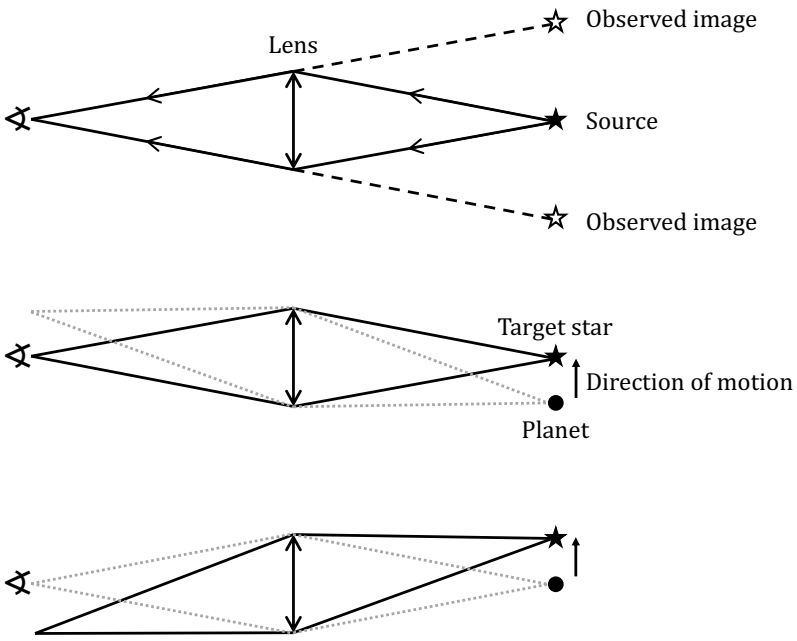

Figure 1.8.: Side view of a simple illustration microlensing event. The light from the source (a distance planetary system) travels to the observer and gets bent by the gravitational lens. Top panel demonstrates the observation of the apparent image, middle and bottom panels demonstrate the change of observable light from the target system. The middle panel can be related to the peak of the microlensing light curve as shown in figure 1.9, while the bottom panel illustrated the second peak that caused by the microlensing event of the planet.

Gaia space observatory operated by the European Space Agency (ESA) was designed to provide high precision astrometric observations (Perryman et al., 2001; Casertano et al., 2008; Perryman et al., 2014). The ongoing Gaia astrometric observations could provide many useful updates for the exoplanetary parameters, such as the inclinations of exoplanet orbits, and a more accurate set of masses of the planets and the stars.

Microlensing

According to the theory of general relativity (GR) by Einstein (1916), gravity bends the light paths through space-time. When a star passes-though the line of sight of the observer and a more distant star, the gravity of the foreground star acts like a lens which increases the apparent brightness (but not the stellar surface brightness) of the source. Figure 1.8 demonstrates a basic illustration case of a mircolensing event where a foreground astronomical object acts as a gravitational lens. The change of alignment between the observer, the lens, and the target system would lead to a

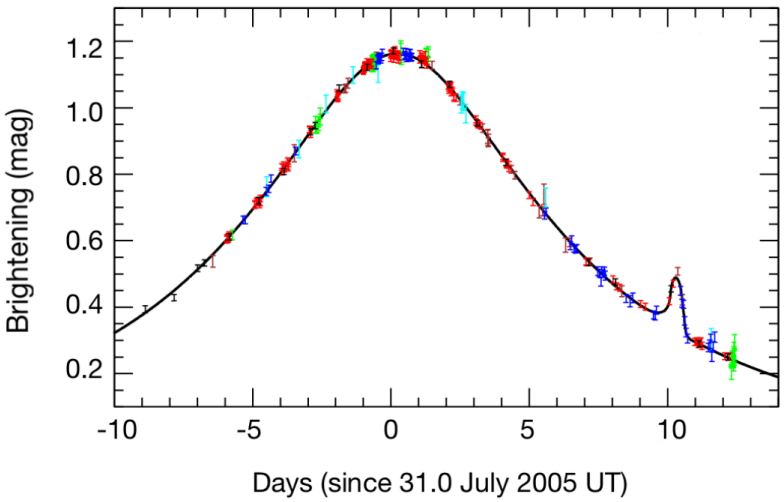

Figure 1.9.: The light curve of the microlensing event of OGLE-20050BLG-390. The highest peak caused by the microlensing event of the host star. A small peak at around 10 days after the maximum peak was caused by the microlensing event of the planet in the system. Data from six different telecospes provides the information in this figure, where orange data points are from La Silla Danish; blue from Perth; black from OGLE; brown from MOA; cyan from Faulkes North; magenta from Tasmania (Beaulieu et al., 2006; Dominik, 2010).

change in apparent brightness of the target system. Figure 1.9 shows an example of an exoplanet microlensing event of PGLE-2005-BLG-390 in 2005 (Beaulieu et al., 2006), where the first peak of the light curve is caused by the microlensing of the star. The second and the smaller peak apeared at around 10 days after the maximum peak, and were caused by the microlensing of the planet.

One of the advantages of this method is being able to detect further out Earths/ super-Earths ($\gtrsim$ 1 au; $\sim$ 1 to 5 $M_\oplus$), such as OGLE-2016-BLG-1195L b (Shvartzvald et al., 2017) and OGLE-2017-BLG-0173L b (Hwang et al., 2018), and point to a different population of exoplanets which is difficult to pick out by the transit and radial velocity methods.

However, a significant defect of the microlensing method is that the detection can not be repeated. This is because of the requirement for a specific alignment between the lens and the target stars. Follow-up detections are also difficult to achieve by other detection methods due to the large scale distance ($\sim$ kpc) between both sources, lens and observer which is required by microlensing.

Optical Gravitational Lensing Experiment (OGLE, Udalski, 2003) and Microlensing Observations in Astrophysics (MOA, Bond et al., 2002) are two of the main

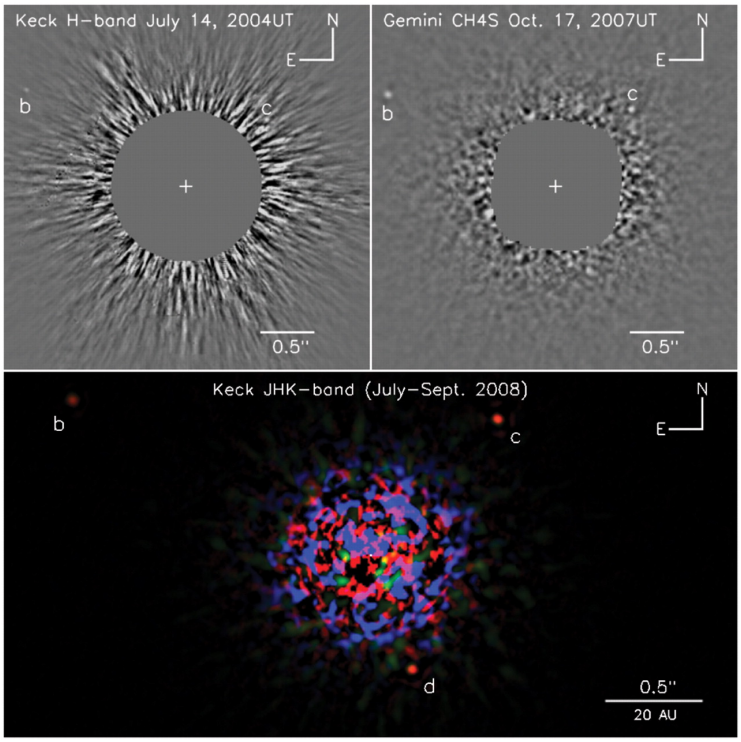

Figure 1.10.: Direct image of system HR 8799, including the three giant planets HR 8799 b, c, and d. The mass of the planet HR 8799 b is ~ 7 M_J. And the planets HR 8799 c and d have a mass of ~ 10 M_J (Marois et al., 2008).

ground-based microlensing observation projects. These two projects are both located in the Southern Hemisphere for better observations of the Galatic bulge. They are located on two sides of the Earth (MOA in New Zealand and OGLE in Chile), which allows the cooperation with each other for continuous observations.

Direct imaging

Unlike many other exoplanet detection methods, direct imaging involves direct observations of the planets. As mentioned earlier in this section, the light from the planet is extremely difficult to detect with today's observational precision. For planets to be observed by direct imaging, they have to be far away from their host star (around a few au to a few thousands au), large enough ($> M_J$), and generally hot enough (to be brighter in the infrared emission). Figure 1.10 shows an example of direct imaging of an exoplanetary system, HR 8799, where the direct images of planets HR 8799 b, c, and d are captured (Marois et al., 2008).

Higher observational precision is required to achieve direct images of exoplanets.

The angular resolution of a telescope, θ_t, can be calculated by

$$\theta_t = 1.22\frac{\lambda}{D_a}, \tag{1.10}$$

where θ_t is in radians, λ is the wavelength of light observed, and D_a is the aperture diameter of the telescope. The relation of $\theta_t \propto 1/D_a$ means that the angular resolution can be improved by doing observations with a larger telescope. Ground-based observations with adaptive optics (Beckers, 1993) and space telescopes (e.g. the Hubble Space Telescope) are commonly used in direct imaging, which can reduce the wavefront distortion by the atmosphere and maintain a high enough resolution for the image.

Orbital brightness modulation

The brightness of a stellar system with a planet that we can observe is not merely based on the star itself. Although the star dominated most of the brightness, a small influence from the planet may still provide some useful information. Transit (section 1.2.2) and reflection are the two examples that planets influence the system's brightness. The light reflected by a planet orbiting a star can provide a periodic change in the total brightness. Like the transit method, the changes of the brightness could also provide evidence that the system might contain planets.

This method has a tendency to detect some relatively large planets or ones with very small orbital radius, and high albedo. This is because the reflection rate from this kind of planets is relatively high. Kepler-70 b and Kepler-70 c (also known as KOI-55 b and c, Charpinet et al., 2011) are the first two examples of confirmed planets discovered by this method. Both of them have an orbital period of less than 0.5 days.

Timing variations

There are gravitational interactions between each object in a planetary system, and some of the interactions might lead to a noticeable variation in their orbital periods. Two main discoveries using this phenomenon are pulsar timing variations and transit timing variations.

As the pulsars have very regular and short rotation periods, the radio waves emitted are also at a highly regular rate. The presence of the planet around it would causes a small movement around the barycentre (similar to figure 1.1), and lead to a small changes in the periodic radio waves observations. The very early

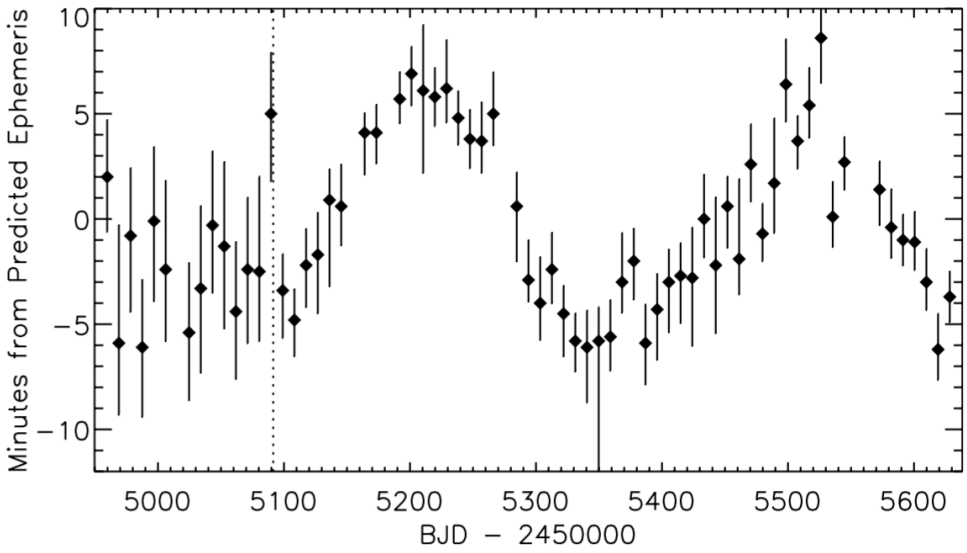

Figure 1.11.: The individual timing deviations of Kepler-19b, which shows a clear variation between the observed and predicted transit time with a period of 316 days and O-C amplitude of 5 minutes (Ballard et al., 2011).

discoveries of exoplanets, PSR B1257+12 c and PSR B1257+12 d were found by the pulsar timing variations observation (Wolszczan and Frail, 1992).

Transit timing variation (TTV) makes use of the periodic change of the timing in the transit event of a planet. This change can be caused by the existence of another non-transiting planet. In the case of the Kepler-19 system, the orbit of the transiting planet, Kepler-19b, had a TTV signal amplitude (maximum time difference between time of predicted transit and observed transit, normally described as observed – calculated, O–C) of $\sim$ 5 minutes. Figure 1.11 shows the time variations (O–C) data of Kepler-19b, which led to the discovery of Kepler-19c (Ballard et al., 2011).

1.3. Exoplanets statistics

1.3.1. Diversity of planetary systems

More than 4000 exoplanets have been discovered by multiple detection methods to date (figure 1.12). The transit method has discovered the greatest number of exoplanets so far (around 76 %), the radial velocity method comes in second and has discovered around 19 % of the confirmed planets, and other methods (e.g. microlensing and direct imaging) contributed the remaining 5 % of discoveries.

Figure 1.13 shows the mass and semi-major axis distributions of the confirmed planets with their discovery detection method. Different types of planets can be

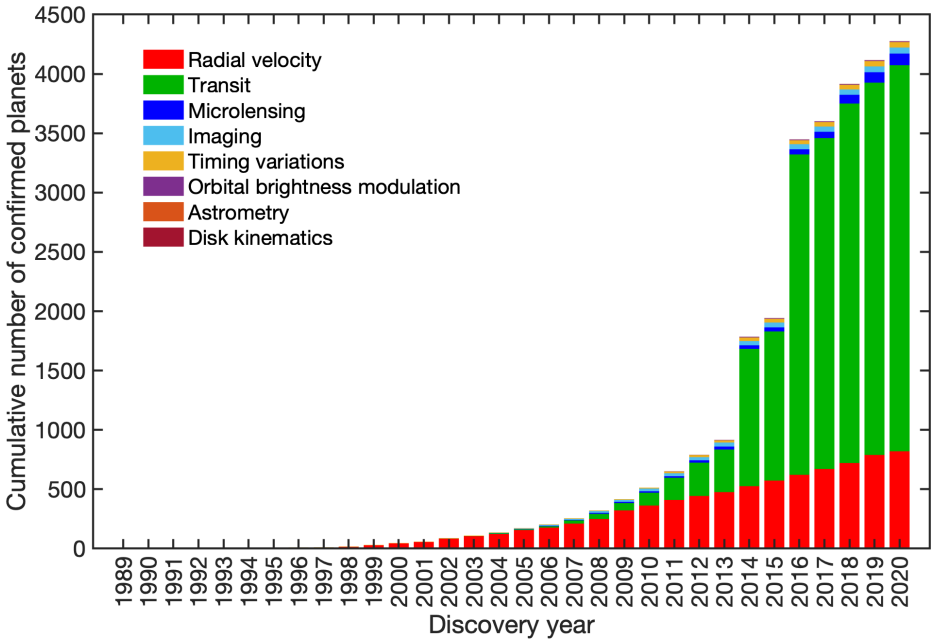

Figure 1.12.: Cumulative number of confirmed exoplanet with their discovery methods from 1989 to September 2020.

discovered by different detection methods (section 1.2). For example, radial velocity observations have a high efficiency to detect Jupiter mass planets with intermediate orbital periods, transit detections are efficacious in detecting planets in a close-in orbit, microlensing can detect bodies at intermediate distances from their host stars, and direct imaging can discover some very large planets orbiting further away from their host star.

Exoplanets can be divided into a few subgroups in terms of their sizes, compositions and orbital distances, for instance, hot-Jupiters, cold gas giants, ice-giants, terrestrial planets, etc. The large population of exoplanets on the top left corner of figure 1.13 is classified as hot-Jupiters, and most of them are discovered by the transit survey or RV observations. A Hot-Jupiter is a gas giant with very short orbital period of $\lesssim$ 10 days ($\lesssim$ 0.1 au), such as some early Kepler discoveries including Kepler-5b (Koch et al., 2010) and 6b (Dunham et al., 2010). Cold giants are classified as the gas giants in more temperate orbits or further out in their system. Microlensing, direct imaging, and radial velocity provided discoveries on a population of cold giants. Direct imaging discovered some cold giants with orbital semi-major axes greater than 10 au, which are difficult to be confirmed by detection methods involving the detection of stellar periodic effects (transit and radial velocity).

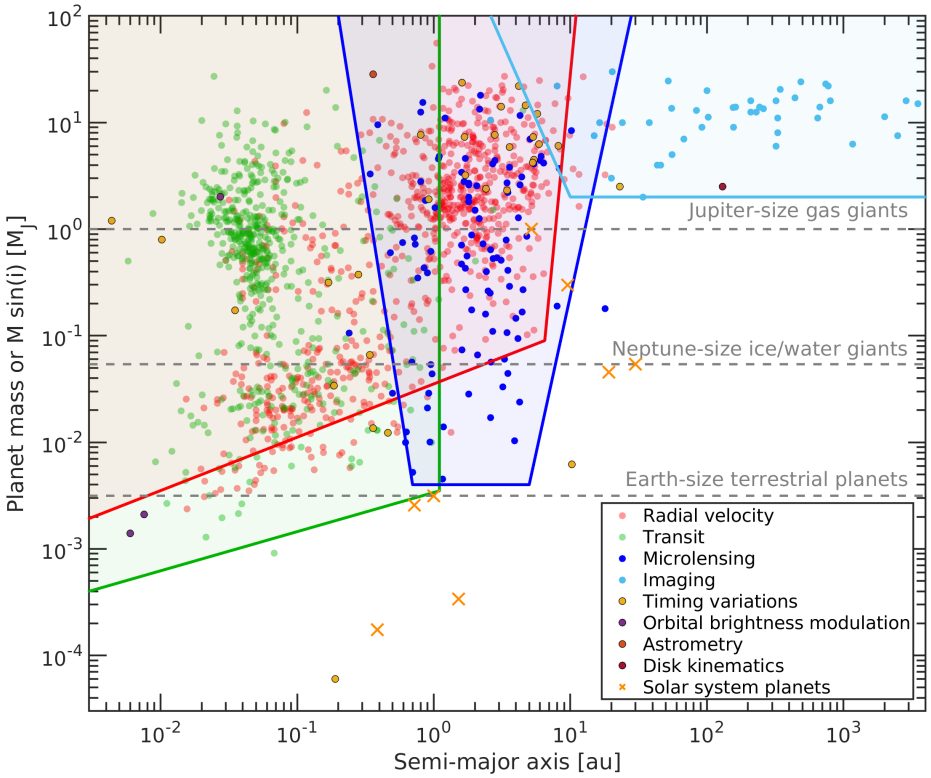

Figure 1.13.: The distribution of the planetary mass (or $M_{\rm p} \sin i_{\rm p}$) and the semi-major axis of confirmed exoplanets. Different colours represent the discovery methods of the relative exoplanets. Solid lines and coloured areas adumbrate the schematic parameter spaces of the radial velocity (red), transit (green), microlensing (blue), and direct imaging (cyan). The horizontal dashed line denoted the mass of Earth (bottom), Neptune (middle), and Jupiter (top).

There are also populations of smaller exoplanets discovered, and they are between the size of Neptune and Earth, namely sub-Neptunes or super-Earths, depending on their sizes and compositions. In general, a super-Earth refers to an exoplanet with a planetary mass higher than that of the Earth, but considerably less massive than Neptune; a sub-Neptune refers to a more massive exoplanet, but still below the mass of Neptune. More specifically, the term super-Earth can be used to describe an exoplanet without a significant atmosphere, while a sub-Neptune has a thick hydrogen-helium atmosphere. These two populations are more close-in (< 1 au), because of the detection methods which are sensitive to the relative sizes (transit and radial velocity) and are not effective in detecting planets with long orbital periods.

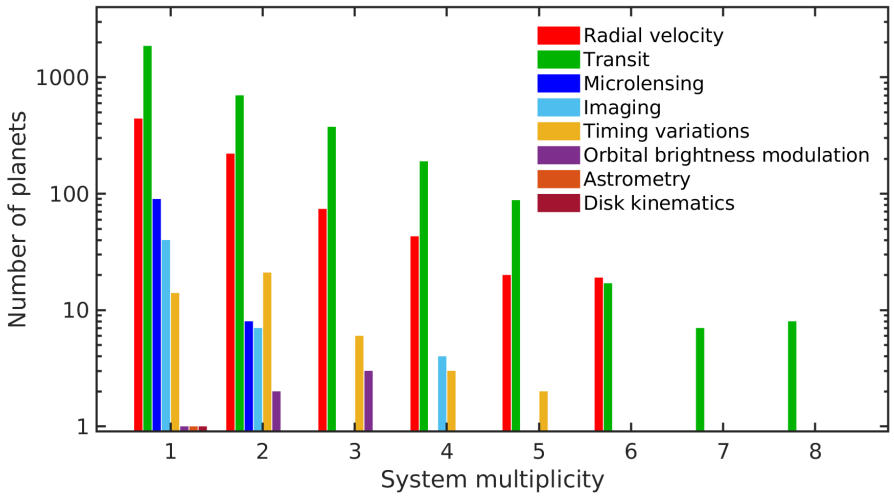

Figure 1.14.: Number of confirmed exoplanets in terms of the multiplicity of their system. The colours represent the discovery method of the planet.

43 % of confirmed planets are found in multiplanet systems (multiplicity ≥ 2). Figure 1.14 shows the number of planets with respect to their system multiplicity. The system multiplicity of Kepler-90 is 8, which makes it the highest multiplicity system that has been discovered so far (Cabrera et al., 2014; Shallue and Vanderburg, 2018). All confirmed planets in the Kepler-90 system have been discovered by the transit method using the Kepler data. The radial velocity method can also detect planetary systems with high multiplicity, such as the HD 10180 (Lovis et al., 2011), HD 219134 (Vogt et al., 2015; Motalebi et al., 2015), and HD 34445 (Howard et al., 2010; Vogt et al., 2017), each of these systems having a multiplicity of 6.

1.3.2. Kepler statistics

More than 55 % of confirmed exoplanets were discovered utilizing Kepler observations. The types of Kepler systems are found to be very diverse, including systems of compact super-Earths and terrestrial planets, single transiting hot Jupiters, and multiplanet systems of terrestrial planets and giants (figure 1.15). The unique proportion of the Kepler discoveries provided a valuable statistical data set, which can lead to better studies about the underlying formation process for the inner parts of planetary systems. Comparing theoretical or simulation results of different planetary system formation hypotheses to the Kepler observation statistic are essential. For example, the multiplicity, eccentricity, and planetary radius distributions from the Kepler observations could be important targets for theoretical studies.

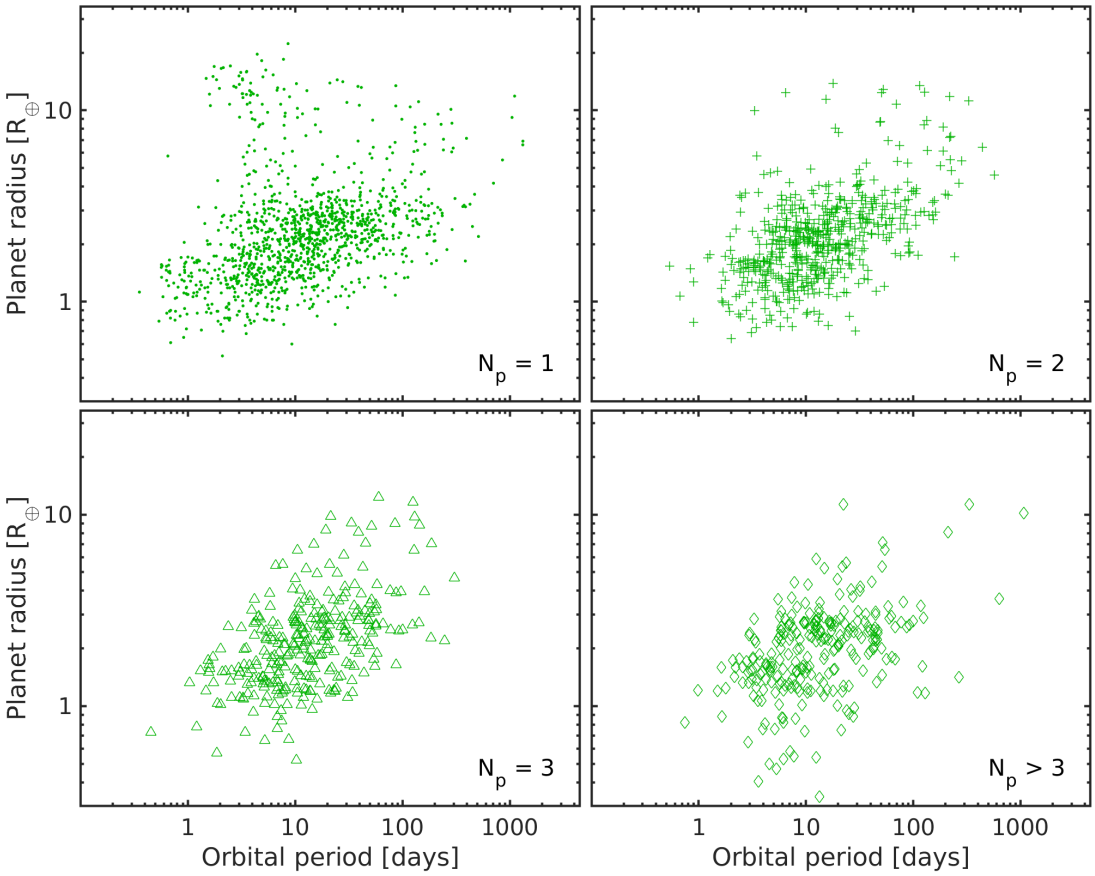

Figure 1.15.: Distributions of confirmed Kepler planets in terms of their planet radii, orbital periods, and mulitiplicities (top left, $N_{\rm p} = 1$; top right, $N_{\rm p} = 2$; bottom left, $N_{\rm p} = 3$; bottom right, $N_{\rm p} > 3$;).

Multiplicity dichotomy

The number of transiting planets or the transit multiplicity of a system is highly dependent on the intrinsic multiplicity, orbit separations between each planet, and mutual inclination of the system. For example, an intrinsic closely packed 5-planet system with a system mutual inclination of 0.001° would have a higher chance to be detected as a multi-transiting system compared to an intrinsic largely spaced 2-planet system with mutual inclination of 5°.

Kepler observations show that there is a large population of single-transit systems relative to the multi-transit systems. Table 1.2 shows the transit multiplicity ratio, TMR(i:j) (define as the number ratio between i-transit systems and j-transit systems), obtained from the Kepler data. It is clear that the TMR(2:1) is smaller than the other TMR by a factor of ~ 1.6. Previous studies show that the TMRs with single-transit systems are difficult to reproduce by models with a single distribution

Table 1.2.: Transit multiplicity ratio (TMR) of the Kepler observation.

Ratio (it:jt)	2t:1t	3t:2t	4t:3t	5t:4t
TMR	0.23	0.36	0.37	0.39

of mutual inclination (Lissauer et al., 2011b; Johansen et al., 2012; Poon, 2016). This implies that the Kepler systems may contain two sub-populations, where one population has high multiplicities and small mutual inclinations, and the other has low multiplicities or large mutual inclinations. The origin of this Kepler multiplicity dichotomy is still not well understood.

Alternatively, Zink et al. (2019) try to explain the multiplicity dichotomy by considering the incompleteness of the Kepler data. Their model shows the reduced detection efficiency after detecting the first planet in a multi-planet system. This detection bias can result in the over-abundance of detected single-transit systems.

Eccentricity dichotomy

Orbital eccentricities of transiting planets can be constrained by comparing their expected and observed transit durations (equations 1.7 and 1.8). The analysis by Xie et al. (2016) first showed that there is an eccentricity dichotomy within the Kepler observations, where the population of Kepler single-transit planets has a higher mean eccentricity, $\langle e_1 \rangle = 0.32$, compared to the population in multi-transit systems, $\langle e_{\geq 2} \rangle = 0.04$. A more recent study by Mills et al. (2019) also shows support for this dichotomy signal with $\langle e_1 \rangle = 0.21$ and $\langle e_{\geq 2} \rangle = 0.05$.

The distribution model of eccentricities commonly follows a Rayleigh distribution (e.g. Zhou et al., 2007; Ford et al., 2008),

$$\text{PDF} = \frac{e}{\sigma_e} \exp\left(-\frac{e^2}{2\sigma_e^2}\right) : 0 \leq e < 1, \tag{1.11}$$

where σ_e is the scale parameter of the eccentricity distribution (eccentricity parameter). Mills et al. (2019) suggested that the values of $\sigma_e = 0.167$ for single-transit systems and $\sigma_e = 0.035$ for multi-transit systems would provide the best agreement with the data (figure 1.16). The strong eccentricity dichotomy signal shown by Xie et al. (2016) and Mills et al. (2019) (see also Van Eylen et al., 2019) indicates the population of single-transit and multi-transit systems could have experienced different dynamical histories.

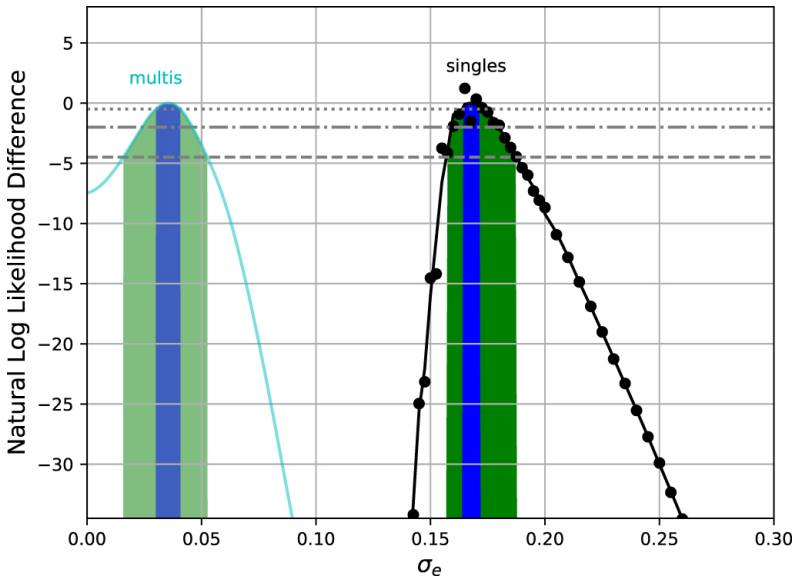

Figure 1.16.: The likelihood of the eccentricity parameter σ_e (Mills et al., 2019). The data and model fit for the eccentricities of single-planet systems are plotted in black. The cyan curve is their model fit for the multi-planet systems. The three horizontal lines indicate the 1- (dotted), 2- (dotted–dashed), and 3- (dashed) σ likelihood differences. The 1-σ confidence intervals are coloured in blue, and 3-σ confidence intervals are coloured in green. This shows a strong dichotomy signal in eccentricity distribution between single-transit and multi-transit systems. It shows that the best fit is with $\sigma_e = 0.167$ for single-planet systems, and $\sigma_e = 0.035$ for multi-planet systems.

Radius gap

The study by Owen and Wu (2013) pointed out that the mass-loss due to photoevaporation of planet atmospheres would lead to a bimodal distribution of the planet size. Later on, the Kepler data analysis by Fulton et al. (2017) provided strong evidence that the radius distribution of small planets ($R_\mathrm{p} < 4$ R$_\oplus$) has separated into two groups (figure 1.17, see also Fulton and Petigura, 2018; Van Eylen et al., 2018). This bimodal distribution shows the two modes at ~ 1.3 R$_\oplus$ and ~ 2.4 R$_\oplus$ and a distribution valley (radius gap) at $1.5 - 2.0$ R$_\oplus$. It indicates that the super-Earth population (~ 1.3 R$_\oplus$) and the sub-Neptune population (~ 2.4 R$_\oplus$) are two distinct classes of planet.

The model of atmosphere photoevaporation suggested that the position of the radius gap can be explained by the planet core composition, where a planet with radius smaller than the position of the valley has a rocky core rather than a volatile rich core (Owen and Wu, 2017; Jin and Mordasini, 2018). The origin of the radius

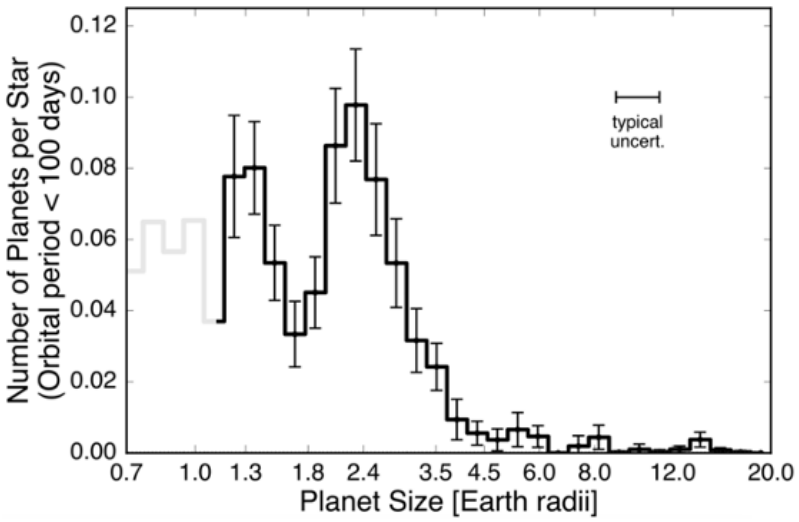

Figure 1.17.: Radius distribution of close-in planets ($P < 100$ days). This shows a bimodal distribution with the modes at ~ 1.3 $R_\oplus$ and ~ 2.4 $R_\oplus$ (Fulton et al., 2017).

gap can also provide information for the planet evolution histories, for example, a study by Owen and Campos Estrada (2020) suggested that a large number of short period super-Earths have lost their gas envelopes by atmosphere photoevaporation.

1.4. Planetary system formation

The two main components of planetary systems are stars and planets. The formation and evolution theories of both components are not completely well understood, and remain as active research areas in astrophysics.

One of the most commonly accepted models of planetary system formation is the nebular hypothesis. The idea was first proposed by Swedenborg (1734) and developed throughout the following century (e.g. Kant, 1755; de Laplace, 1796). The modern model for this hypothesis is the solar nebular disc model (SNDM) by Safronov (1972), which applied only to the Solar system at first, but then was suggested to apply to other planetary systems. The minimum mass solar nebula (MMSN, Hayashi, 1981) is the solar disc model which gives the minimum mass of solid material for the formation of the 8 Solar system planets.

The nebular hypothesis indicates that the stars are formed in giant molecular clouds. These giant molecular clouds are gravitationally unstable. Young protostars are then formed by the rotation and gravitational collapse of the clouds. If the

Figure 1.18.: 5 protostars in the Orion Nebula, where 4 of the protostars are surrounded by their protostellar discs (credit to C. R. O'Dell, ESA / Hubble; released on 13-6-1994).

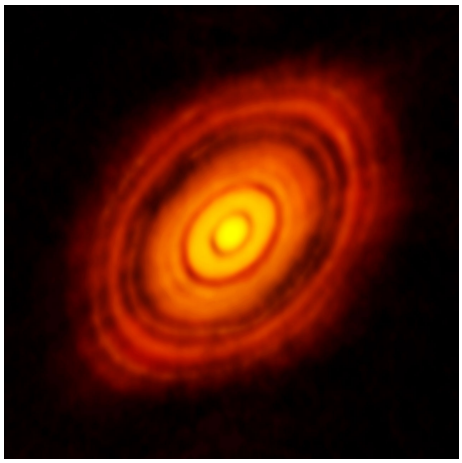

Figure 1.19.: A protoplanetary disc surrounding HL Tauri. Image taken by Atacama Large Millimeter Array (ALMA, ALMA Partnership et al., 2015).

angular momentum of the cloud is greater than zero (rotating), a protostellar disc can be formed (figure 1.18). The materials which are further away from the protostar with higher angular momentum would orbit around the newly formed star, in the form of a protoplanetary disc (figure 1.19).

The disc has a rotational motion around the central body, and we can use that to define the orbital plane. The upward and downward moving particles would cancel out their vertical momentum due to the conservation of momentum. So, the gas and dust particles above or below the orbital plane tend to move closer to the orbital plane, and the disc becomes flat. At this stage, dust grains in the flat disc can continue to merge together to form bigger objects, such as planetesimals, protoplanets, and planets, until the disc dissipates.

1.4.1. Accretion disc

A protoplanetary disc is an accretion disc that supplies material to the central star. This thin disc generally has negligible mass compared to the host star, where the typical disc mass is around 1 to 3 % of the stellar mass. The central mass domination provides a good approximation that the angular velocity of the disc complies with Kepler's 3rd Law (equation 1.5), therefore a Keplerian disc. The disc material moves in nearly circular orbits, with a slow inward drift (accretion). This is due to the angular momentum transfer mechanisms. Outward transport of angular momentum occurs, which allows disc material to lose angular momentum and drift inwards.

Viscosity is considered to be one possible angular momentum transport mechanism. The Keplerian orbital angular velocity has an inverse relation to the orbital radius. The inner material rotates faster than the outer material, which induces a friction force such that the inner material loses angular momentum to the outer material.

The thickness of a viscous disc is supported by hydrostatic equilibrium. If we assume the gas is locally isothermal (constant local temperature from the disc midplane to surface), the half-thickness, H, of the disc can be calculated by

$$H^2 = \frac{\mathcal{R}T}{\mu\Omega^2}, \tag{1.12}$$

where $\mathcal{R}$ is the ideal gas constant, T is the local temperature, μ is the molecular mass, and Ω is the local Keplerian velocity. The local Keplerian velocity at orbital distance, r is $\Omega = \sqrt{GM_\star/r^3}$. The isothermal sound speed, c_s, in an ideal gas is calculated by

$$c_s = \sqrt{\frac{\mathcal{R}T}{\mu}}. \tag{1.13}$$

Combining equations 1.12 and 1.13, a relation between the local sound speed, Keplerian velocity, and half-thickness can be written as

$$c_\mathrm{s} = H\Omega, \tag{1.14}$$

which means that the thickness of the Keplerian disc can be computed once the molecular mass and local isothermal temperature at each distance r is known.

The viscosity arises from the frictional force between gas at different orbital distances which orbit at different speeds. Consider an annulus of disc material at distance r from the centre with a width $\mathrm{d}r$, in a Keplerian disc; the time evolution

of the local surface density due to viscosity is governed by (Pringle, 1981),

$$\frac{\partial \Sigma}{\partial t} = \frac{1}{r}\frac{\partial}{\partial r}\left[3r^{1/2}\frac{\partial}{\partial r}\left(\nu \Sigma r^{1/2}\right)\right], \tag{1.15}$$

where t is the time, ν is the kinematic viscosity, and Σ is the local surface density of the disc.

1.4.2. Terrestrial planet formation

A terrestrial planet, or rocky planet, is a planet that has a rocky core which dominates the total mass of the planet. Mercury, Venus, Earth, and Mars are examples of rocky planets in the Solar system. In general, the idea of the formation process for such planets is believed to follow several key stages.

Small dust grains tend to settle towards the mid-plane of the protoplanetary disc. The settling time-scale, $\tau_{\rm set}$, of the grains can be approximated as,

$$\tau_{\rm set} \approx \frac{3\Sigma}{16\pi R_{\rm dg}\rho_{\rm dg}} \text{ orbits}, \tag{1.16}$$

where $R_{\rm dg}$ is the radius of the dust grain, and $\rho_{\rm dg}$ is the density of the dust grain. A typical settling time-scale is around 10^5 yrs at 5 au. The agglomeration of dust grains during this period of time occurs via direct collisions, and builds up centimetre-scale particles.

Kilometre-sized planetesimals may be formed by further coagulation between those centimetre sized particles (Safronov, 1972; Wetherill and Stewart, 1989). The time-scale of this process might be similar to $\tau_{\rm set}$. On the other hand, Youdin and Goodman (2005) suggested an alternative mechanism where planetesimals may form via gravitational collapse of groups of larger particles, known as the streaming instabilities (see also, Johansen et al., 2007).

The large planetesimals continue their growth towards protoplanets via direct collisions and accretion of smaller background objects, where the factor of gravitational focusing becomes important. A large planetesimal can gravitationally focus other bodies towards it leading to a bigger collision cross-section than its physical cross-section. By the conservation of energy and conservation of angular momentum, the collision cross-sections can be given by

$$\pi b^2 = \pi R_{\rm pt}^2\left[1 + \left(\frac{v_{\rm esc}}{v_r}\right)^2\right], \tag{1.17}$$

where b is the impact parameter (also the effective radius of collision of the growing protoplanet), $R_{\rm pt}$ is the physical radius of the protoplanet, $v_{\rm esc}$ is the escape velocity of the protoplanet, and v_r is the relative velocity of the incoming object. This yields the mass accretion rate,

$$\frac{{\rm d}M_{\rm pt}}{{\rm d}t} = nm_b v_r \pi R_{\rm pt}^2 \left[1 + \left(\frac{v_{\rm esc}}{v_r}\right)^2\right], \tag{1.18}$$

where $M_{\rm pt}$ is the mass of the protoplanet, n is the number density, and m_b is the mass of the background object.

When the escape velocity is much larger than the background velocity dispersion, $v_r \ll v_{\rm esc}$, the protoplanet would be considered in the 'runaway growth' regime. The protoplanet in this regime can grow very rapidly and dissociate from the background mass distribution. The rate of change of the protoplanet mass during the runaway growth is proportional to the mass with a power index of 4/3 (${\rm d}M_{\rm pt}/{\rm d}t \propto M_{\rm pt}^{4/3}$ or $\propto R_{\rm pt}^4$), which indicates that the growth rate is larger when the mass is bigger (runaway). Typically, the time-scale of runaway growth of a protoplanet is $\sim 10^5$ yrs. This runaway growth would be slowed down when the protoplanet has consumed most of the planetesimals within its reach or becomes massive enough to increase the velocities of the background planetesimals. When the velocity dispersion of the background planetesimals is comparable to the escape velocity, $v_r \simeq v_{\rm esc}$, the radius grows linearly with time ($R_{\rm pt} \propto t$). In the case where $v_r \gg v_{\rm esc}$, gravitational focusing factor is less important and this results in slow growth. This growth rate is much slower than the runaway growth rate but still faster than the growth of planetesimals. This regime is referred as 'oligarchic growth' (Kokubo and Ida, 1998).

The protoplanets can grow by accreting surrounding planetesimals until their isolation mass is approached. The isolation mass, $M_{\rm iso}$, is given by

$$M_{\rm iso} = 4\pi r \Sigma_{\rm s} \mathcal{K} R_{\rm Hill,pt}, \tag{1.19}$$

where $\Sigma_{\rm s}$ is the surface density of the solid, $\mathcal{K}$ is a constant, and $R_{\rm Hill,pt}$ is the Hill radius of the protoplanet at orbital distance r. This is obtained by the protoplanet which has accreted all of the planetesimals within an annulus of width $2\mathcal{K}R_{\rm Hill,pt}$ (the feeding zone). For a MMSN, the isolation mass in the Earth's feeding zone is around 6 lunar masses and roughly from a few to 10 Earth masses in Jupiter's feeding zone. The low isolation mass obtained suggests that the terrestrial planets in the solar system did not stop their growth after accreting only planetesimals within their feeding zones. Further evolutions can be caused by collisions and mergers

between the protoplanets.

Besides the planetesimal accretion mechanism discussed above, recent studies proposed that protoplanets can also grow by accretion of mm to cm-sized pebbles (Johansen and Lacerda, 2010; Ormel and Klahr, 2010). In this pebble accretion scenario, the drag force applied to the pebbles from the surrounding disc gas can be important. When a pebble is inside the Hill/Bondi sphere of a protoplanet, the pebble would spiral onto the protoplanet due to the loss of angular momentum via the drag force (see also Lambrechts and Johansen, 2012; Morbidelli and Nesvorny, 2012; Ida et al., 2016; Visser and Ormel, 2016; Johansen and Lambrechts, 2017; Liu and Ji, 2020). The growth of protoplanets under this mechanism is not limited by the disc material within the feeding zone. Pebble accretion can continue until there is a lack of disc gas around the protoplanets (e.g. gas disc dissipation or the protoplanets are massive enough for a gap opening). Thus, the final mass of a protoplanet depends more on the stellar and disc mass, instead of the local environment of the disc in the planetesimal accretion mechanism. In general, the final core mass obtained from pebble accretion is around 10 to 20 $M_\oplus$ for a Sun-like star.

1.4.3. Giant planet formation

The large amount of gas, mainly hydrogen and helium, contained in gas giants implies that the giants must have formed within the protoplanetary disc lifetime, which is around a few million years. Two scenarios for the formation process have been proposed and studied extensively: the core accretion model (e.g. Perri and Cameron, 1974; Mizuno et al., 1978; Bodenheimer and Pollack, 1986; Pollack et al., 1996) and the gravitational instability model (e.g. Cameron, 1978; Boss, 1997).

Core accretion model

The core accretion model is the evolution process that continued after the rocky core formation discussed in the previous section (section 1.4.2). When a rocky core becomes massive enough, its strong gravitational potential can hold on to a bound and significant gas envelope. The gas envelope accretion can be triggered when the surface escape velocity of the evolving planet is larger than the local sound speed in the protoplanetary disc. Initially, this gas accretion rate is very slow, but accelerates with time.

For a massive core ($\gtrsim$ 3 $M_\oplus$), the gas accretion may lead to an envelope that is massive enough, and no longer maintains hydrostatic equilibrium. It will then undergo the Kelvin-Helmholtz contraction, followed by a stage of runaway gas ac-

cretion. The runaway gas accretion can increase the planet envelope mass rapidly and become dynamically significant to its planet-planet interactions. This rapid process will be slowed down when the planet is massive enough to open a gap in the protoplanetary disc, which reduces the rate of gas accretion.

For a low-mass rocky core, this gas accretion process is very gentle. The time-scale for such a core to reach the runaway accretion regime would be longer than the disc lifetime, resulting in an atmosphere that is dynamically insignificant.

The core accretion model hits an obstacle when explaining the formation of widely separated giants. Direct imaging observations (section 1.2.3) pointed to a population of giants which are orbiting far away from their hosts ($\gtrsim$ 10 au). At such orbital distance, the growth of planet cores is not effective enough to form cores that can follow the core accretion model to become giants. However, massive cores which are greater than 10 $M_\oplus$ can still form in the later stages of the disc lifetime; those cores would not be able to have enough time for gas accretion to reach the runaway gas accretion phase. For such populations of giants, alternative models may be required to explain the formation histories.

Gravitational instability model

Alternatively, a portion of giant planets might form under the gravitational instability mechanism. The main idea of gravitational instability is from the fragmentation of a protoplanetary disc. If the disc is very massive (comparable to the stellar mass), it may be gravitationally unstable and lead to fragmentation. Under this circumstance, gravitational collapse of such fragments can lead to giant planet formation.

The condition for such instabilities to occur requires a small value of the Toomre Q_T parameter (Toomre, 1964), which is

$$Q_T = \frac{c_s \Omega}{\pi G \Sigma} \lesssim 1. \tag{1.20}$$

This is the essential requirement, but not a sufficient condition for fragmentation to happen. For example, the instabilities can generate density waves (spiral arms) that can heat up the disc, which increases the local sound speed as well as the value of Q_T, and make fragmentation more difficult to happen. This suggests that a relatively short cooling time is important for disc fragmentation to happen (e.g. Gammie, 2001, suggested that the cooling time has to be $\lesssim 3\Omega^{-1}$).

2. Methods and models

打開你的宇宙 便萬有
錯與對一念如魔成佛
愛恨有盡時即管參透
努力去面對是美麗以後珍惜所有

— 方皓玟 〈唵嘛呢叭咪吽〉

N-body simulations are a powerful tool to undertake studies of planetary system evolution. Planetary N-body simulation codes offer a fast and reliable platform to achieve accurate simulations of the orbital motions of bodies due to their gravitational interactions. For example, in the case of our Solar System, a 4.6 billion years old and long-term stable system, these N-body algorithms can provide an excellent prediction of its future. However, to study the history and the formation process when the planetary system is young, when the protoplanetary disc still existed and/or giant impacts of protoplanets were common, the traditional N-body algorithms would face difficulty providing a comprehensive simulation. To achieve a more realistic simulation result, I adopted different physical models for the N-body integration, including a realistic collision model and a protoplanetary disc model. In this chapter I describe the general numerical model that I adopted in my N-body simulations. The specific sets of physical models adopted in different studies are presented in later chapters. Moreover, I describe the synthetic transit observation technique which can provide synthetic transit surveys of planetary systems from simulations, like a virtual Kepler telescope, in this chapter.

2.1. Collision model

Traditional N-body simulation codes, such as MERCURY (Chambers, 1999) and SYMBA (Duncan et al., 1998), assume all planet-planet collisions result in perfect mergers (figure 2.1). It is a reasonable assumption for planet-planet collisions that happen at large semimajor axis and low eccentricity, where the impact energy of the

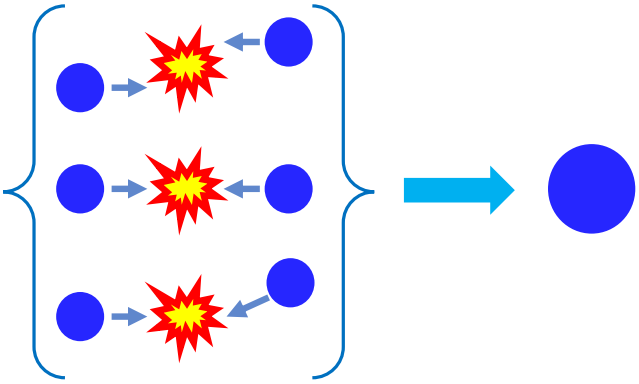

Figure 2.1.: Traditional perfect accretion in N-body simulations. All collision events end up with perfect mergers.

collision event is low. For a close-in planetary system, the giant impacts can occur at high eccentricity. The high impact energy can result in significant fragmentation of the collided bodies. In such collision events, instead of considering only perfect mergers, the collision outcome would be more diverse (figure 2.2). To achieve a more practical simulation model, I implemented a 'realistic' prescription into the N-body code to handle the collision outcome[1]. This realistic collision model is based on the prescription of Leinhardt and Stewart (2012); I refer the reader to that paper for the detailed description. In the following sections, I describe the collision condition parameters that are relevant to my adopted collision model, and discuss the collision outcomes that applied to my N-body simulation.

2.1.1. Collision parameters

For the two bodies involved in the collision, I refer to the more massive object as the target and the less massive object as the projectile. The total mass involved in the collision event is simply the sum of the mass of the two colliding bodies, $M_{\rm Total} = M_1 + M_2$, where subscripts 1 refer to the target and 2 refer to the projectile ($M_1 \geq M_2$). The specific impact energy with respect to the centre of mass is calculated according to

$$Q_{\rm R} = \frac{1}{2}\frac{\mu}{M_{\rm Total}} V_{\rm imp}^2, \tag{2.1}$$

where $\mu = (M_1 M_2)/(M_1 + M_2)$ is the reduced mass, and $V_{\rm imp}$ is the relative impact velocity.

[1]I wish to thank Alessandro Morbidelli and Seth Jacobson for providing their version of the SYMBA N-body code, which included the prescriptions of defining the realistic collision regimes (section 2.1.2) and collision debris generation (section 2.1.3).

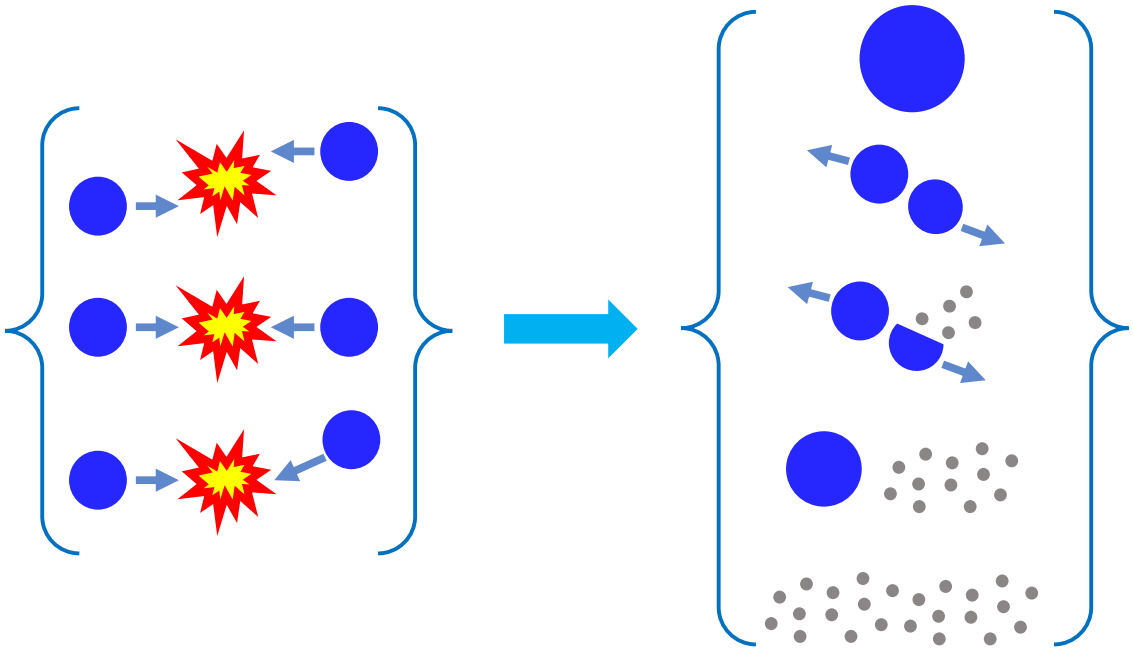

Figure 2.2.: Realistic accretion adopted in the N-body simulation. Different collision conditions could end up with different collision outcomes.

The value of Q_{R} also indicates the specific impact energy for a head-on collision. For an off-centre collision, a correction to Q_{R} is suggested by Leinhardt and Stewart (2012). Figure 2.3 demonstrates a geometric illustration of an off-centre collision, where θ is the impact angle, R_1 is the radius of the target, R_2 is the radius of the projectile, and l is the projected length of the projectile overlapping the target. The projected length can be calculated by

$$l = \begin{cases} R_{\mathrm{Total}}(1-b), & \text{if } b \geq \dfrac{R_1 - R_2}{R_{\mathrm{Total}}} \\ 2R_2, & \text{otherwise (entire projectile interacts)} \end{cases} \tag{2.2}$$

where $R_{\mathrm{Total}} = R_1 + R_2$ is the total radius, and $b = \sin\theta$ is the impact parameter. The part of the projectile which is outside the region of the projected length does not directly intersect the target, different from a head-on collision. For an oblique collision with a large impact angle, the collision is likely to be a grazing impact. The critical impact parameter for a grazing impact is $b_{\mathrm{crit}} = R_1/(R_1 + R_2)$, where this critical value indicates that half of the projectile directly interacts with the target (Asphaug, 2010).

For the off-centre collision correction, we have to consider not only the whole part of the collision system, but also the part that is directly involved in the collision.

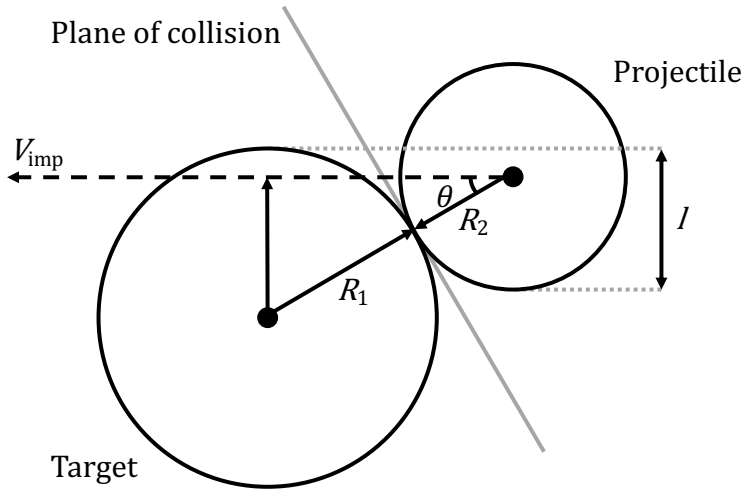

Figure 2.3.: Geometric illustration of an off-centre collision and the definitions of different collision parameters.

The total interacting mass is defined as $M_\alpha = M_1 + \alpha M_2$, where α is the fraction of mass of the projectile that directly interacts with the target. Following the geometric model in figure 2.3, α can be calculated by

$$\alpha = \frac{M_{2,\text{interact}}}{M_2} = \frac{3R_2 l^2 - l^3}{4R_2^3} = \frac{3R_2 R_{\text{Total}}^2 (1-b)^2}{4R_2^3}, \tag{2.3}$$

where $M_{2,\text{interact}}$ is the mass of the projectile that is estimated to be directly involved in the collision. The interacting reduced mass between the target and the interacting part of the projectile can then be calculated by, $\mu_\alpha = (\alpha M_1 M_2)/(M_1 + \alpha M_2)$. Similarly, the escape velocity of the whole collision system is $V_{\text{esc}} = \sqrt{2GM_{\text{Total}}/R_{\text{Total}}}$, and the interacting escape velocity can be calculated by,

$$V_{\text{esc},\alpha} = \sqrt{\frac{2GM_\alpha}{R_{\text{Total}}}}. \tag{2.4}$$

Following the prescription in Leinhardt and Stewart (2012) for the off-centre collision, the specific impact energy to disperse half the total mass in an oblique impact, Q'^*_{RD}, can be calculated by

$$Q'^*_{\text{RD}} = -\frac{4}{5} c^* \pi \rho_1 G R_{C1} \left(\frac{\mu}{\mu_\alpha}\right)^{2-3\bar{\mu}/2} \left[\frac{1}{4}\left(1 + \frac{M_2}{M_1}\right)^2 \frac{M_1}{M_2}\right]^{2/(3\bar{\mu})-1}, \tag{2.5}$$

where c^* is a dimensionless material parameter which describes the disruption energy of a head-on and equal-mass collision in terms of the gravitational binding energy, $\bar{\mu}$

is the coupling parameter of the energy and momentum of the projectile (Holsapple and Schmidt, 1987), and R_{C1} is calculated by

$$R_{C1} = \sqrt[3]{\frac{3M_{\rm Total}}{4\pi\rho_1}}, \tag{2.6}$$

which is the radius of total mass in a body with density, $\rho_1 = 1$ g cm^{-3} (Stewart and Leinhardt, 2009). The hydrodynamic simulation results in Leinhardt and Stewart (2012) give a best-fit value of $c^* = 5.0$ and $\bar{\mu} = 0.37$ in the small body regime ($R_{C1} \lesssim 0.5$ R$_\oplus$), and $c^* = 1.9$ and $\bar{\mu} = 0.36$ in the planet size regime ($R_{C1} > 0.5$ R$_\oplus$). With the calculated value of $Q'^*_{\rm RD}$, the critical impact velocity, V'^*, is given by

$$V'^* = \sqrt{\frac{2Q'^*_{\rm RD}M_{\rm Total}}{\mu}}. \tag{2.7}$$

This critical impact velocity can then be directly compared to the value of $V_{\rm imp}$ obtained from the N-body simulation and distinguish the types of collision outcome.

In addition, there are some collision events that are considered in the 'reverse impact scenario', where only a fraction of the target is directly involved in the impact with the projectile. In the reverse case, the whole projectile is involved in the collision ($\alpha = 1$). The involved mass from the target, $M^\dagger$, yields the total interacting mass, $M^\dagger_{\rm Total} = M_2 + M^\dagger$, and the reverse reduced mass, $\mu^\dagger = (M_2 M^\dagger)/(M_2 + M^\dagger)$. The specific reverse impact energy, $Q^\dagger_{\rm R}$, is given by

$$Q^\dagger_{\rm R} = \frac{1}{2}\frac{\mu^\dagger}{M^\dagger_{\rm Total}}V^2_{\rm imp}, \tag{2.8}$$

and the reverse radius of the total mass in the body with ρ_1, $R^\dagger_{C1}$, follows the relation of $R_{C1}(M_{\rm Total}) \rightarrow R^\dagger_{C1}(M^\dagger_{\rm Total})$ (equation 2.6). This leads to the critical reverse impact energy, $Q^{\dagger*}_{\rm RD}$, to be

$$Q^{\dagger*}_{\rm RD} = -\frac{4}{5}c^*\pi\rho_1 G R^{\dagger}_{C1}\left(\frac{\mu^\dagger}{\mu_\alpha}\right)^{2-3\bar{\mu}/2}\left[\frac{1}{4}\left(1+\frac{M^\dagger}{M_2}\right)^2\frac{M_2}{M^\dagger}\right]^{2/(3\bar{\mu})-1}. \tag{2.9}$$

We can define different collision regimes by applying different values of the critical impact energy, $Q'_{\rm R,regime}$, or the critical reverse impact energy, $Q^\dagger_{\rm R,regime}$. Similar to

equation 2.7, the critical impact velocity is given by

$$V'_{\rm regime} = \sqrt{\frac{2Q'_{\rm R,regime}M_{\rm Total}}{\mu}}, \tag{2.10}$$

for the normal impact scenario, or

$$V^{\dagger}_{\rm regime} = \sqrt{\frac{2Q^{\dagger}_{\rm R,regime}M^{\dagger}_{\rm Total}}{\mu^{\dagger}}}, \tag{2.11}$$

for the reverse impact scenario, where different collision regimes and outcomes are discussed in the following section (section 2.1.2).

2.1.2. Collision outcomes

In all the collision events, the total mass before and after the collisions is always conserved, and maintains the relation of

$$M_{\rm Total} = M_{\rm LR} + M_{\rm SLR} + M_{\rm T,d}, \tag{2.12}$$

where $M_{\rm LR}$ is the mass of the post-collision's largest remnant, $M_{\rm SLR}$ is the second largest remnant, and $M_{\rm T,d}$ is the total mass of the debris particles that are generated during the collision.

The collision outcomes of our adopted realistic collision model are defined in nine different regimes: Perfect Merger, Super-catastrophic, Catastrophic, Erosion, Partial Accretion, Hit-and-Spray, Hit-and-Run, Bounce, and Graze-and-Merge Regime. Depending on the collision condition and the collision parameters that are obtained in the previous section (section 2.1.1), each collision event would be assigned to one of the nine regimes. Figure 2.4 shows the layout of the decision tree for the collision algorithm adopted in the N-body simulation.

Perfect merger regime

This is the only regime that is assumed in the traditional N-body simulation. In my collision model, collisions result in this regime when

$$V_{\rm imp} < V_{{\rm esc},\alpha}. \tag{2.13}$$

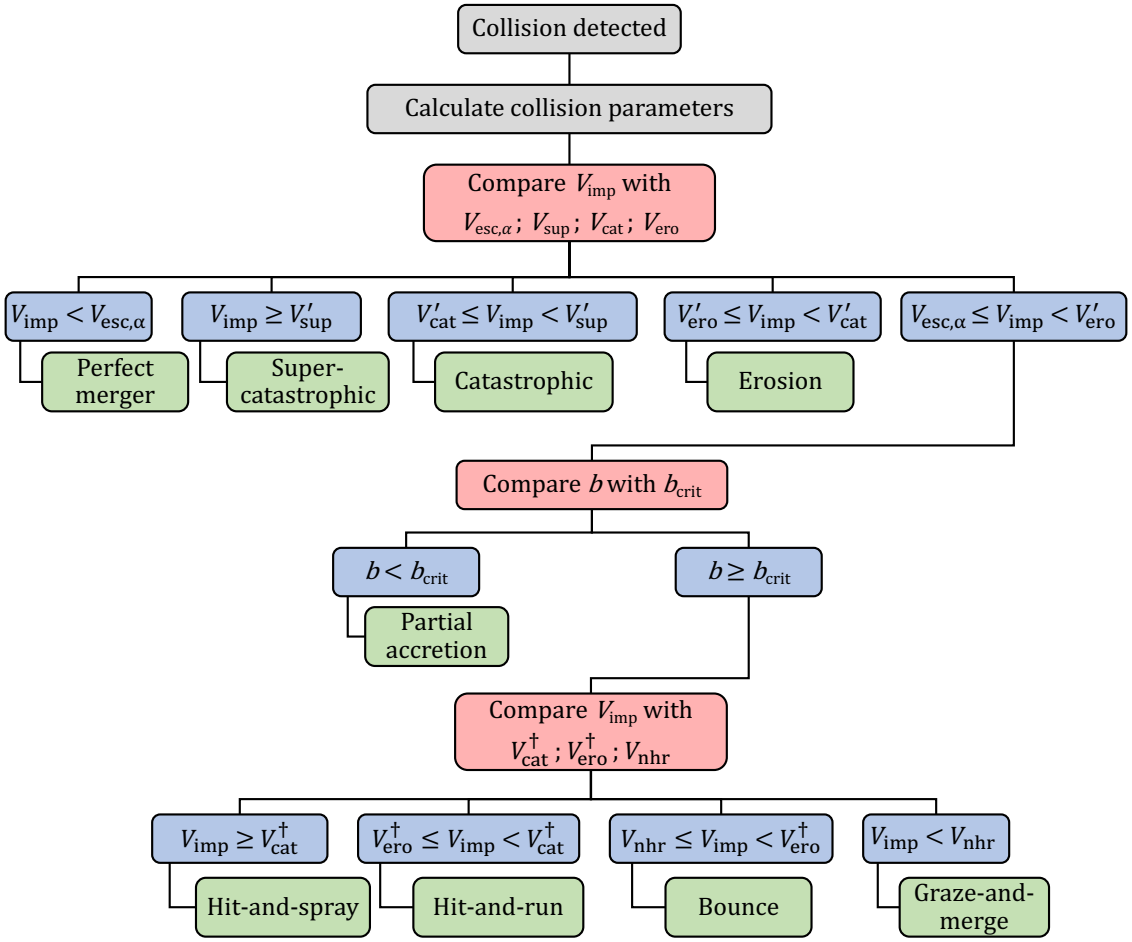

Figure 2.4.: Decision tree of the collision outcome that is adopted in the N-body simulation.

As a result, a single body is formed with the total mass and momentum of the original two collided bodies, where

$$M_{\rm LR} = M_{\rm Total}; \qquad M_{\rm SLR} = 0; \qquad M_{\rm T,d} = 0. \tag{2.14}$$

Super-catastrophic regime

This is the most energetic collision regime within all the possible outcomes in the collision model. Collisions are considered in this regime when

$$V_{\rm imp} \geqslant V'_{\rm sup}, \tag{2.15}$$

where $V'_{\rm sup}$ is the critical impact velocity for super-catastrophic collision. The value of $V'_{\rm sup}$ is given by

$$V'_{\rm sup} = \sqrt{\frac{2Q'_{\rm R,sup}M_{\rm Total}}{\mu}} \quad , \text{ with} \quad Q'_{\rm R,sup} = \frac{76}{39}Q'^{*}_{\rm RD}. \tag{2.16}$$

In this regime, the impact energy is so high that both the target and projectile break into debris, as

$$M_{\rm LR} = 0; \qquad M_{\rm SLR} = 0; \qquad M_{\rm T,d} = M_{\rm Total}. \tag{2.17}$$

Catastrophic regime

The catastrophic regime is considered when

$$V'_{\rm cat} \leqslant V_{\rm imp} < V'_{\rm sup}, \tag{2.18}$$

where the $V'_{\rm cat}$ is the threshold velocity for a catastrophic collision, which is given by

$$V'_{\rm cat} = \sqrt{\frac{2Q'_{\rm R,cat}M_{\rm Total}}{\mu}} \qquad \text{with} \qquad Q'_{\rm R,cat} = \frac{9}{5}Q'^{*}_{\rm RD}. \tag{2.19}$$

The high impact energy results in only one massive body remaining and the rest of the mass is in the form of collision debris, following the relation of

$$M_{\rm LR} = 0.1M_{\rm Total}\left(\frac{5}{9}\frac{Q_{\rm R}}{Q'^{*}_{\rm RD}}\right)^{-3/2}; \qquad M_{\rm SLR} = 0; \qquad M_{\rm T,d} = M_{\rm Total} - M_{\rm LR}. \tag{2.20}$$

Erosion regime

An erosion collision happens when

$$V'_{\rm ero} \leqslant V_{\rm imp} < V'_{\rm cat}, \tag{2.21}$$

where $V'_{\rm ero}$ is the threshold of the erosion regime. The value of $V'_{\rm ero}$ is given by

$$V'_{\rm ero} = \sqrt{\frac{2Q'_{\rm R,ero}M_{\rm Total}}{\mu}} \qquad \text{where} \qquad Q'_{\rm R,ero} = 2Q'^{*}_{\rm RD}\frac{M_2}{M_1}\left(1 + \frac{M_2}{M_1}\right)^{-1}. \tag{2.22}$$

In an erosion collision, the target is eroded by the projectile and loses some of its mass, while the projectile is completely disrupted into a number of collision particles,

following

$$M_{\rm LR} = M_{\rm Total}\left(1 - \frac{Q_{\rm R}}{2Q'^{*}_{\rm RD}}\right); \qquad M_{\rm SLR} = 0; \qquad M_{\rm T,d} = M_{\rm Total} - M_{\rm LR}. \tag{2.23}$$

Partial accretion regime

Partial accretion happens when

$$b < b_{\rm crit} \qquad \text{and} \qquad V_{{\rm esc},\alpha} \leqslant V_{\rm imp} < V'_{\rm ero}. \tag{2.24}$$

The impact energy is below the erosion threshold but above the perfect merger threshold, and additionally has a low impact angle. The post-collision mass distribution follows the relation stated in equation 2.23. In this regime, the target gains some mass from the projectile and the remaining part of the projectile is completely disrupted into collision debris.

Hit-and-spray regime

Hereafter, the collision models enter the high impact angle regime. The collision condition for a hit-and-spray collision to happen is

$$b \geqslant b_{\rm crit} \qquad \text{and} \qquad V_{\rm imp} \geqslant V^{\dagger}_{\rm cat}, \tag{2.25}$$

where $V^{\dagger}_{\rm cat}$ is the reverse impact velocity threshold of a hit-and-spray collision, which is calculated by

$$V^{\dagger}_{\rm cat} = \sqrt{\frac{2Q^{\dagger}_{\rm R,cat}M^{\dagger}_{\rm Total}}{\mu^{\dagger}}} \qquad \text{where} \qquad Q^{\dagger}_{\rm R,cat} = \frac{9}{5}Q^{\dagger *}_{\rm RD}. \tag{2.26}$$

This collision allows the target to retain its mass, while the projectile is completely disrupted into debris particles, following

$$M_{\rm LR} = M_1; \qquad M_{\rm SLR} = 0; \qquad M_{\rm T,d} = M_2. \tag{2.27}$$

Hit-and-run regime

The impact condition for a hit-and-run collision is

$$b \geqslant b_{\rm crit} \qquad \text{and} \qquad V^{\dagger}_{\rm ero} \leqslant V_{\rm imp} < V^{\dagger}_{\rm cat}, \tag{2.28}$$

where $V^{\dagger}_{\rm ero}$ is the threshold of the impact velocity required, where

$$V^{\dagger}_{\rm ero} = \sqrt{\frac{2Q^{\dagger}_{\rm R,ero}M^{\dagger}_{\rm Total}}{\mu^{\dagger}}} \qquad \text{with} \qquad Q'_{\rm R,ero} = 2Q^{\dagger *}_{\rm RD}\frac{M^{\dagger}}{M_2}\left(1+\frac{M^{\dagger}}{M_2}\right)^{-1}. \tag{2.29}$$

In the case of a hit-and-run collision, the target can retain its mass and the projectile loses some of its mass in the form of collision debris, as

$$M_{\rm LR} = M_1; \quad M_{\rm SLR} = M^{\dagger}_{\rm Total}\left(1-\frac{Q^{\dagger}_{\rm R}}{2Q^{\dagger *}_{\rm RD}}\right); \quad M_{\rm T,d} = M_{\rm Total} - M_{\rm SLR} - M_1. \tag{2.30}$$

Note that the hit-and-run regime defined here is different from the hit-and-run collision that is considered in the studies by Kokubo and Genda (2010). The hit-and-run collision in Kokubo and Genda (2010) has no mass lost in the form of collision debris, and this type of collision is considered as 'bounce collision' in my model.

Bounce regime

A bounce collision happens when

$$b \geqslant b_{\rm crit} \qquad \text{and} \qquad V_{\rm hnr} \leqslant V_{\rm imp} < V^{\dagger}_{\rm ero}, \tag{2.31}$$

where $V_{\rm hnr}$ is the threshold of a bounce collision. The value of $V_{\rm hnr}$ follows the relation given in Kokubo and Genda (2010) and Genda et al. (2012), which is

$$V_{\rm hnr} = \left[C_1\Gamma^2\left(1-b\right)^{C_5} + C_2\Gamma^2 + C_3\left(1-b\right)^{C_5} + C_4\right]V_{\rm esc}, \tag{2.32}$$

where $\Gamma = (M_1 - M_2)/M_{\rm Total}$, and C_1 to C_5 are constants with the value of $C_1 = 2.43$, $C_2 = -0.0408$, $C_3 = 1.86$, $C_4 = 1.08$, and $C_5 = 5/2$. In the bounce collision regime, no mass is lost in the form of debris and both bodies retain their masses:

$$M_{\rm LR} = M_1; \qquad M_{\rm SLR} = M_2; \qquad M_{\rm T,d} = 0. \tag{2.33}$$

Graze-and-merge regime

A graze-and-merge collision happens when the impact velocity is low but the impact angle is large:

$$b \geqslant b_{\rm crit} \qquad \text{and} \qquad V_{\rm imp} < V_{\rm hnr}. \tag{2.34}$$

The collision outcome of this regime is similar to the perfect merger, as in equation 2.14; one object remains which contains all the mass of the two colliding bodies.

2.1.3. Collision debris

In my collision model, collision debris can be produced in a super-catastrophic, catastrophic, erosion, partial accretion, hit-and-spray, or hit-and-run collision. With the known total mass of the debris that is obtained in equations 2.17, 2.20, 2.23, 2.27, or 2.30, the parameters of the debris are allocated in terms of the total number of particles, $N_{\rm d}$, position vector, $\mathbf{r}_{\rm d}$, and velocity vector, $\mathbf{V}_{\rm d}$.

The number of debris is allocated by

$$N_{\rm d} = \begin{cases} \max\left[\dfrac{M_{\rm T,d}}{M_{\rm d,max}}, \min\left(38, \dfrac{M_{\rm T,d}}{M_{\rm d,min}}\right)\right], & \text{if } M_{\rm T,d} > 0 \\ 0, & \text{otherwise} \end{cases} \tag{2.35}$$

where $M_{\rm d,max}$ and $M_{\rm d,min}$ are the maximum and minimum mass of the debris, which are tunable simulation parameters. If the value of $N_{\rm d}$ is not an even integer, it is rounded-up to the nearest even integer. The total mass is then evenly distributed to each debris particle.

The debris particles produced are evenly distributed in a circle on the plane of impact, with their velocity vectors pointed away from the collision centre of mass (figure 2.5). The positions of the debris are distributed according to

$$\mathbf{r}_{\rm d} = \mathbf{r}_{\rm com} + \epsilon R_{\rm Hill} \hat{\mathbf{r}}_{\rm d}, \tag{2.36}$$

where $\mathbf{r}_{\rm com}$ is the position vector of the collision centre of mass, $\hat{\mathbf{r}}_{\rm d}$ is the position unit vector with respect to the collision centre of mass, $R_{\rm Hill}$ is the Hill radius of the collision system, and ϵ is a constant that can be adjusted to an appropriate value for the simulation. Similarly, the velocity vector for the debris is

$$\mathbf{V}_{\rm d} = \mathbf{V}_{\rm com} + \varepsilon V_{\rm esc} \hat{\mathbf{r}}_{\rm d}, \tag{2.37}$$

where $\mathbf{V}_{\rm com}$ is the velocity vector of the collision centre of mass, and ε is a constant that can be adjusted to an appropriate value for the simulation. In general, ε has a value of $\varepsilon > 1$ to avoid quick re-accretion between the debris and the two largest remnants.

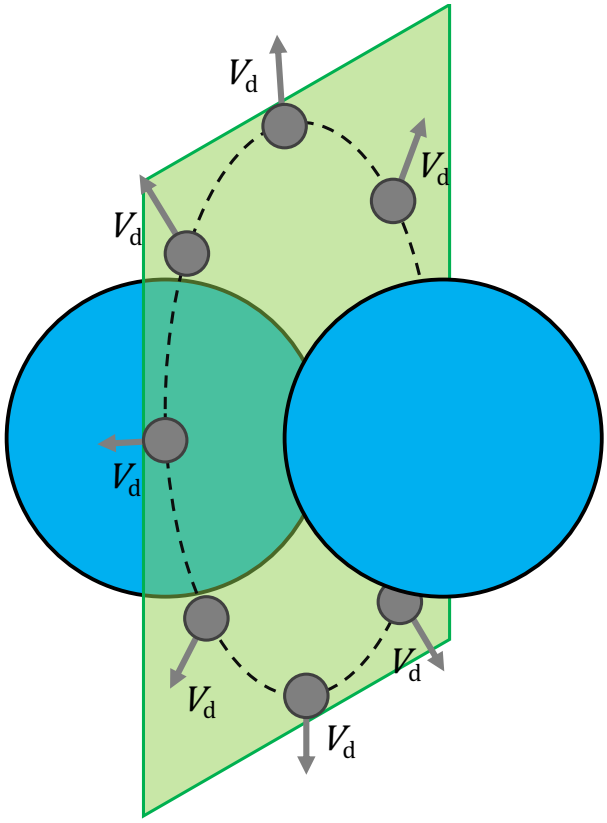

Figure 2.5.: Illustration of a realistic planet-planet collision. The two blue circles are the colliding target and projectile, the green plane is the plane of collision, and the grey circles are the collision debris produced. Collision debris are located around a circle with radius $\epsilon R_{\rm Hill}$ (equation 2.36) on the plane of collision (dashed line). The initial velocity of each debris (grey arrows) has a value of $\varepsilon V_{\rm esc}$ (equation 2.37) with respect to the collision centre of mass.

2.2. Protoplanetary disc model

To provide N-body simulations for planetary systems in an earlier phase, before the protoplanetary disc has dissipated, I incorporated a disc model into the N-body code on top of the realistic collision model. The basic physical model and numerical methods are described in this section.

2.2.1. Initial disc profile

The two key initial parameters that are required in this disc model are the initial surface density profile, $\Sigma_{\rm init}(r)$, and the temperature profile, $T(r)$, where r is the distance from the centre object. We follow the power index for the minimum mass solar nebula (MMSN: Hayashi, 1981) to set up our initial condition of the gas disc.

The initial surface density profile is obtained by

$$\Sigma_{\rm init}\left(r\right) = \Sigma_{\rm init}\left(1{\rm au}\right) r^{-1.5}, \tag{2.38}$$

where $\Sigma_{\rm init}\left(1{\rm au}\right)$ is the initial surface density of the disc at the distance of 1 au. The value of $\Sigma_{\rm init}\left(1{\rm au}\right)$ is adjustable for different situations. The temperature profile also follows the power index of the MMSN and is calculated by

$$T\left(r\right) = T\left(1{\rm au}\right) r^{-0.5}, \tag{2.39}$$

where $T\left(1{\rm au}\right)$ is the temperature of the disc at 1 au. In my disc model, the temperature is assumed to retained for the whole simulation before the gas disc dissipated. The value of $\Sigma_{\rm init}\left(1{\rm au}\right)$ is set as 1700 g cm^{-2} and $T\left(1{\rm au}\right)$ is equal to 280 K in the studies contained in this thesis unless stated otherwise.

2.2.2. Viscous disc evolution

I use the α model for the kinematic viscosity in our viscous disc model, namely the α-disc model. With the dimensionless viscosity parameter, α, the kinematic viscosity is given by

$$\nu(r) = \alpha c_{\rm s} H = \alpha \Omega H^2 = \alpha \frac{c_{\rm s}^2}{\Omega}, \tag{2.40}$$

by connecting $c_{\rm s}$, H, and Ω from equation 1.14. The value of α is set as 1×10^{-3} in the studies contained in this thesis unless stated otherwise.

A 1-dimensional disc evolution is adopted in the simulation model. The fundamental equation that governs the disc evolution, similar to equation 1.15, follows the diffusion equation of the disc surface density,

$$\frac{\partial \Sigma}{\partial t} = \frac{1}{r}\frac{\partial}{\partial r}\left[3r^{1/2}\frac{\partial}{\partial r}\left(\nu\Sigma r^{1/2}\right) - \frac{2\Lambda\Sigma r^{3/2}}{\left(GM_\star\right)^{1/2}}\right], \tag{2.41}$$

where Λ is the injection rate of angular momentum per unit mass by tidal interaction with the protoplanets in the disc (equation 5.3, see also Lin and Papaloizou, 1986). The extra term including Λ on the right of equation 2.41 represents the angular momentum exchange from the protoplanets that acted on the disc. For a protoplanet which is massive enough, this tidal interaction term would result in a gap opening in the disc model (e.g. equation 5.4).

Finite Difference Equations

To numerically solve the diffusion equation that governs the disc surface density evolution (equation 2.41), a finite differencing scheme is applied. The numerical grid cells of the disc is defined by the distance from the central star, following the radial profile of

$$r(i) = r_{\rm in}\left(\frac{r_{\rm out}}{r_{\rm in}}\right)^{(i-1)/(n-1)}, \tag{2.42}$$

where $r_{\rm in}$ is the inner boundary of the disc, $r_{\rm out}$ is the outer boundary of the disc, and i denotes the i-th grid cell out of n total grid cells. The radial profile adopted allows a high resolution in the inner region of the disc, and can benefit the studies that focused on the close-in planetary systems.

The centred finite difference expression for the viscous term, $\nu\Sigma r^{1/2}$, in equation 2.41 is expressed in the form of

$$\frac{\partial}{\partial r}\left(\nu\Sigma r^{1/2}\right) = \frac{\left(\nu\Sigma r^{1/2}\right)_{i+1} - \left(\nu\Sigma r^{1/2}\right)_{i-1}}{\Delta r_{i-1} + \Delta r_i}, \tag{2.43}$$

where Δr_i is the physical width of the i-th grid cell. Analogous to the first derivative, the centred finite difference expression for the second derivative is given by

$$\frac{\partial^2}{\partial r^2}\left(\nu\Sigma r^{1/2}\right) = \frac{\left(\nu\Sigma r^{1/2}\right)_{i+1} - \left(\nu\Sigma r^{1/2}\right)_{i}}{\Delta r_{i-1}\Delta r_i} - \frac{\left(\nu\Sigma r^{1/2}\right)_{i} - \left(\nu\Sigma r^{1/2}\right)_{i-1}}{(\Delta r_{i-1})^2}. \tag{2.44}$$

The finite difference expression for the planet-disc interaction term, $\Lambda\Sigma r^{3/2}$, is expressed in either forward or backward difference, which depends on the local radial distance and the position of the protoplanet that is interacting with the disc:

$$\frac{\partial}{\partial r}\left(\Lambda\Sigma r^{3/2}\right) = \begin{cases} \dfrac{\left(\Lambda\Sigma r^{3/2}\right)_{i+1} - \left(\Lambda\Sigma r^{3/2}\right)_{i}}{\Delta r_i} & \text{for } v_{\rm rad} < 0, \\ \dfrac{\left(\Lambda\Sigma r^{3/2}\right)_{i} - \left(\Lambda\Sigma r^{3/2}\right)_{i-1}}{\Delta r_{i-1}} & \text{otherwise,} \end{cases} \tag{2.45}$$

where $v_{\rm rad}$ is the local velocity of the disc material travelling in the radial direction of the disc, and is given by

$$v_{\rm rad} = 2\Lambda\left(\frac{r}{GM_\star}\right)^{1/2} - \frac{3}{\Sigma r^{1/2}}\frac{\partial}{\partial r}\left(\nu\Sigma r^{1/2}\right) \tag{2.46}$$

at the radial distance r. This approach was found to improve the stability of the

integration.

Inserting the finite difference of the viscous and planet-disc interaction terms (equation 2.43, 2.44, and 2.45) to the diffusion equation of the disc surface density (equation 2.41), the local surface density of the next time step, Σ_{t+1}, is calculated by

$$\Sigma_{t+1} = \Sigma_t + \Delta t \left[\frac{3}{2r^{3/2}} \frac{\partial}{\partial r} \left(\nu \Sigma r^{1/2} \right) + \frac{3}{r^{1/2}} \frac{\partial^2}{\partial r^2} \left(\nu \Sigma r^{1/2} \right) - \frac{2}{r (G M_\star)^{1/2}} \frac{\partial}{\partial r} \left(\Lambda \Sigma r^{3/2} \right) \right], \tag{2.47}$$

where Σ_t is the surface density at the current time step, and Δt is the size of the disc time step. The value of Δt is taken to be $\Delta t = 0.1 \times \min \left(\Delta r_i / v_{\rm rad} \right)$, which avoids disc material travelling through more than one grid cell radially within a single time step. In general, the disc time step differs from the N-body time step, where a reasonable N-body time step is set to be 1/20th of the shortest orbital period of the protoplanets in the system (Duncan et al., 1998).

In a realistic situation, the protplanetary disc might not only be governed by the viscous evolution and planet-disc interaction, but also by other physical processes, such as photoevaporation that is induced by ultraviolet radiation absorption by the disc. If the additional evolutionary equation of the surface density, $\partial \Sigma_{\rm add} / \partial t$, is known, the additional effect can be simply added to equation 2.41 by

$$\frac{\partial \Sigma}{\partial t} = \frac{\partial \Sigma_{\rm vis}}{\partial t} + \frac{\partial \Sigma_\Lambda}{\partial t} + \frac{\partial \Sigma_{\rm add}}{\partial t}, \tag{2.48}$$

and can be solved numerically by

$$\Sigma_{t+1} = \Sigma_t + \Delta t \left[\frac{\Delta \Sigma_{\rm vis,t}}{\Delta t} + \frac{\Delta \Sigma_{\Lambda,{\rm t}}}{\Delta t} + \frac{\Delta \Sigma_{\rm add,t}}{\Delta t} \right]. \tag{2.49}$$

Boundary conditions

In the numerical method described above, which is applied to solve the disc diffusion equation, the finite difference cannot provide solutions at the boundary. For the outer boundary, I assumed a 'zero-velocity' condition, where the radial velocity at the outer edge of the disc is 0. This assumption means no extra material is added to the disc from the outer edge. The surface density of the outermost grid cell is given by

$$\Sigma_n = \Sigma_{n-1} \frac{\nu_{n-1}}{\nu_n} \left(\frac{r_{n-1}}{r_n} \right)^{1/2}. \tag{2.50}$$

The surface density for the innermost grid cell can be calculated by the net move-

ment of the disc material. The assumption for this calculation is based on the notion that the material added to the innermost grid cell must only come from the second-innermost grid cell, and the material displaced from the inner edge of the disc is accreted into the central star. Under this assumption, the surface density evolution of the innermost grid cell is given by

$$\Sigma_{t+1,1} = \Sigma_{t,1} - \frac{\Delta t}{\Delta r_1}\left(v_{\mathrm{rad},2}\Sigma_{t,2}\frac{r_2}{r_1} - v_{\mathrm{rad},1}\Sigma_{t,1}\right), \tag{2.51}$$

where the radial velocity of the first grid cell, $v_{\mathrm{rad},1}$, is assumed as

$$v_{\mathrm{rad},1} = \min\left(2v_{\mathrm{rad},2} - v_{\mathrm{rad},3}, -\frac{3}{2}\frac{\nu_1}{r_1}\right). \tag{2.52}$$

The first term inside the bracket of equation 2.51 represents the material moved into the innermost grid cell from the second-innermost grid cell, and the second term represent the material accreted into the central star from the innermost cell.

2.2.3. Planet evolution

For an object orbiting in the disc environment, the interaction between the object and the disc is important. The planet-to-disc interaction, such as the angular momentum exchange term (equation 5.3), is discussed in the later section. On the other hand, the disc-to-planet interaction is also taken into account, including the gravitational interaction and the evolution of the planetary structure.

The gravitational interaction between the disc and the orbiting object will induce a perturbation on the orbit of the body. The interaction results in a near-circular orbit for the object, due to eccentricity damping (e.g. Goldreich and Tremaine, 1980; Artymowicz, 1993). Meanwhile, inclination damping from the disc would result in the orbital plane of the body tending to the mid-plane of the disc. Following the expression given in Papaloizou and Larwood (2000), the acceleration on each body is

$$\boldsymbol{a}_{\mathrm{damp}} = \frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t} = -2\frac{(\boldsymbol{v}\cdot\boldsymbol{r})\,\boldsymbol{r}}{t_{\mathrm{e}}r^2} - 2\frac{(\boldsymbol{v}\cdot\boldsymbol{z})\,\boldsymbol{z}}{t_{\mathrm{i}}z^2}, \tag{2.53}$$

where $\boldsymbol{v}$ is the velocity vector of the body, $\boldsymbol{r}$ is the position vector of the body on the disc plane, $\boldsymbol{z}$ is the position vector of the body perpendicular to the disc plane, t_{e} is the damping time-scale of the eccentricity, and t_{i} is the damping time-scale of the inclination.

For a body with mass below the gap opening criterion, eccentricity and inclination

damping are induced by the excitation of the density waves (e.g. Ward, 1988; Tanaka et al., 2002), including in regions close to the planet, which is not applicable for a body that has opened a gap in the disc. In the studies included in this thesis, the damping time-scale for a gap opening object is set to be 20 local orbital periods for both t_e and t_i unless stated otherwise.

The planet formation model applied in my simulation is an *in situ* model, and planetary migration such as type I and type II migration, which causes large scale changes of the orbital semimajor axis, is not considered. More details of the damping mechanism included in the simulation model will be presented in section 5.2.2.

According to the core accretion model (section 1.4.3), a massive protoplanet in the disc can accrete a gas envelope from the surrounding gas. A gas envelope accretion mechanism is also included my simulation model, and it will be discussed in detail in section 5.2.3.

2.3. Synthetic transit observation

Transit surveys have discovered many of the planetary systems with high multiplicities (section 1.2.2). However, it is very likely that any transit survey could not detect all the planets that constitute the systems. The relative number of one to eight-planet systems discovered via transit detections are highly dependent on both the intrinsic multiplicities of the systems, and the mutual orbital inclinations of the systems. Therefore, conducting synthetic transit observations of the simulated planetary systems would allow us to build up a synthetic dataset, which would be meaningful to compare to the actual transit data, such as the one carried out by Kepler (section 1.3.2).

2.3.1. Near-circular orbits

Most of the transiting planets in multiplanetary systems have near-circular orbits (section 1.3.2), and assuming a circular orbit can still provide a good estimation to the Kepler statistical properties (Moorhead et al., 2011). The benefit of assuming a circular orbit is that the transit determination only depends on the orbital plane and the orbital radius, instead of the argument of periapsis.

The orbital plane of the planet can be defined by the angular momentum in the

form of

$$\hat{\mathbf{h}} = \begin{pmatrix} \hat{h}_x \\ \hat{h}_y \\ \hat{h}_z \end{pmatrix} = \begin{pmatrix} \sin I \sin \Omega \\ -\sin I \cos \Omega \\ \cos I \end{pmatrix}, \tag{2.54}$$

where $\hat{\mathbf{h}}$ is the unit vector of the angular momentum, $\hat{h}_x$-$\hat{h}_y$-$\hat{h}_z$ are the x-y-z components of $\hat{\mathbf{h}}$ in Cartesian coordinates, I is the inclination of the planetary orbital plane, and Ω is the longitude of ascending node of the orbital plane. In the N-body simulations, the planetary evolution always follow a fixed reference direction. To adopt different viewing angles to the simulated systems, a coordinate transformation between the viewing position and the system reference plane is applied to every synthetic observation. The system reference plane is defined by the innermost planet, where I introduce a new set of reference Cartesian coordinate in the x'-y'-z' directions. The transformation (figure 2.6) is governed by

$$\begin{pmatrix} \hat{h}_x \\ \hat{h}_y \\ \hat{h}_z \end{pmatrix} = \mathbf{R}_z(\Omega)\mathbf{R}_x(I) \begin{pmatrix} \hat{h}'_x \\ \hat{h}'_y \\ \hat{h}'_z \end{pmatrix}, \tag{2.55}$$

where $\mathbf{R}_x(I)$ and $\mathbf{R}_z(\Omega)$ are the 3×3 rotation matrices along the x-axis and z-axis respectively with elements

$$\mathbf{R}_x(I) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos I & -\sin I \\ 0 & \sin I & \cos I \end{pmatrix}, \tag{2.56}$$

and

$$\mathbf{R}_z(\Omega) = \begin{pmatrix} \cos \Omega & -\sin \Omega & 0 \\ \sin \Omega & \cos \Omega & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{2.57}$$

With the known values of the orbital radius and orbital plane from the simulation, and assuming a circular orbit, the orbit of the planet is fixed. In geometry, a circle in a 3-dimensional space projects on a 2-dimensional plane as an ellipse. Equally, a circular orbit can be projected on different reference planes in the form of an ellipse. The semi-major axis of the projected ellipse has the same value as the circular orbit radius a. Meanwhile, the point of the projected semi-minor axis is always the point to the nearest star surface (in the 2-dimensional plane). From the direction of the angular momentum, we are able to get the relationship to the projected semi-minor

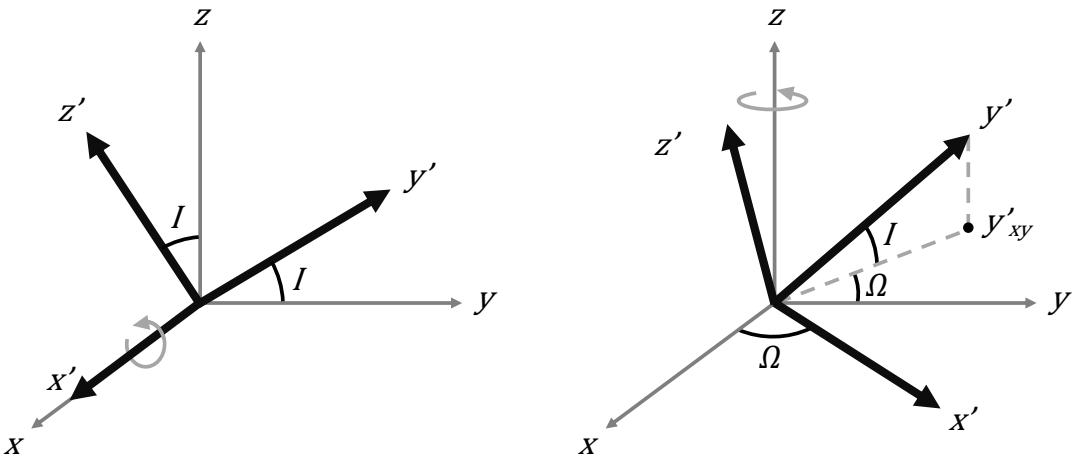

Figure 2.6.: Rotation of axes between the viewing position (x-y-z) and the system reference plane (x'-y'-z'). On the left is the rotation by $\mathbf{R}_x(I)$ where the coordinate system rotate along the x-axis by angle I, and on the right is the rotation by $\mathbf{R}_z(\Omega)$ along the z-axis by angle Ω.

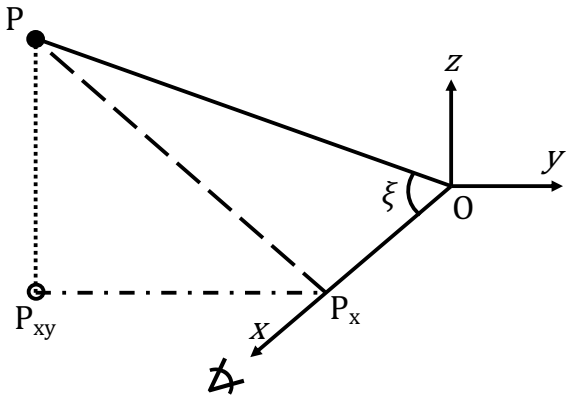

Figure 2.7.: Geometic demonstration of the definition of the angle of angular momentum, ξ, to the x-axis (observing direction).

axis, b_p. In figure 2.7, the length of line OP is 1, and has the same direction as $\hat{\mathbf{h}}$, where P_{xy} is the projection of P on the x-y plane and P_x is the projection of P_{xy} on the x-axis. ξ is defined as the angle between the observer's line of sight (along the x-axis) and angular momentum (marked as $\angle POP_x$ in figure 2.7). A simple trigonometric relationship can be used to find ξ, as the length of $\hat{\mathbf{h}}$ is always 1, which is

$$\cos\xi = \frac{OP_x}{OP} = \frac{\hat{h}_x}{|\hat{\mathbf{h}}|} = \frac{\hat{h}_x}{1} = \sin I \sin\Omega, \tag{2.58}$$

where $\hat{h}_x = \sin I \sin\Omega$ from equation 2.54.

The value of ξ is calculated by equation 2.58 and the length of b_p can be found by

$$b_p = |a \cos \xi|. \tag{2.59}$$

The occurrence of a transit event observed by the synthetic observation can simply be determined by comparing the radius of the star, $R_\star$, and b_p. I consider the orbit to be a transiting orbit when the semi-minor axis of the projected ellipse is overlapping the star. The orbit is transiting the star when

$$b_p \leq R_\star. \tag{2.60}$$

2.3.2. Eccentric orbits

In some cases, when the orbits of the planets are eccentric, the synthetic observation method that is mentioned in the previous section (section 2.3.1) could not provide good accuracy. In an eccentric orbit, the argument of periapsis, ω, is also an important factor to determine the occurrence of the transit orbit.

Instead of the angular momentum, it is more convenient to operate directly on the orbital plane for an eccentric orbit, because of the changing distance between the orbital focus and the planet throughout an orbit. The coordinate transformation of the orbit is governed by (Murray and Dermott, 1999):

$$\hat{\mathbf{p}} = \begin{pmatrix} \cos\Omega\cos(\omega+f) - \sin\Omega\sin(\omega+f)\cos I \\ \sin\Omega\cos(\omega+f) + \cos\Omega\sin(\omega+f)\cos I \\ \sin(\omega+f)\sin I \end{pmatrix}, \tag{2.61}$$

where the true anomaly, f, can be obtained by

$$\tan\frac{f}{2} = \sqrt{\frac{1+e}{1-e}}\tan\frac{E}{2}, \tag{2.62}$$

and E is the eccentric anomaly. $\hat{\mathbf{p}}$ is different throughout the whole orbit and needs to be calculated individually for different values of E, where $0 \leq E < 2\pi$.

Next, by comparing $\hat{\mathbf{p}}$ to a randomly chosen observation direction, $\hat{\mathbf{o}}$, we can define an observation parameter, $\delta_{\min}$, where

$$\delta_{\min} = \min\left(|\hat{\mathbf{p}} - \hat{\mathbf{o}}|\right). \tag{2.63}$$

Similar to equations 2.59 and 2.60, the orbit transits the star when

$$r \sin \delta_{\min} \leqslant R_\star, \tag{2.64}$$

where $r = a(1 - e \cos E)$ is the distance between the planet and the orbital focus at a given position E.

3. Compact super-Earths systems formation

在宇宙旅程 由引力决定
還有種貪玩脱軌 出了軌
往后都不配安定
路線被固定
誰擴大過自己的直徑
代價竟高得震驚
為尋覓火花銷毀了生命

— 王菀之, 林夕 ⟨我來自火星⟩

This chapter mainly follows the material that is contained in Poon et al. (2020). The title of this publication is '*Formation of compact systems of super-Earths via dynamical instabilities and giant impacts*', and it was published in volume 491 of the Monthly Notices of the Royal Astronomical Society.

NASA's Kepler mission discovered $\sim$ 700 planets in multiplanet systems containing three or more transiting bodies, many of which are super-Earths and sub-Neptunes in compact configurations whose origins are not yet understood. Using N-body simulations, we examine the *in situ*, final stage assembly of multiplanet systems via the collisional accretion of protoplanets. Our initial conditions are constructed using a subset of the Kepler five-planet systems as templates, and apply to the epoch after gas disc dispersal. Two different prescriptions for treating planetary collisions are adopted. The simulations address numerous questions: Do the results depend on the accretion prescription?; do the resulting systems resemble the Kepler systems, and do they reproduce the observed distribution of planetary multiplicities when synthetically observed?; do collisions lead to significant modification of protoplanet compositions, or to stripping of gaseous envelopes?; do the eccentricity distributions agree with those inferred for the Kepler planets?

We find that the accretion prescription is unimportant in determining the out-

comes. The final planetary systems look broadly similar to the Kepler templates adopted, but the observed distributions of planetary multiplicities or eccentricities are not reproduced, because scattering does not excite the systems sufficiently. In addition, we find that ~ 1 per cent of our final systems contain a co-orbital planet pair in horseshoe or tadpole orbits. Post-processing the collision outcomes suggests that they would not significantly change the ice fractions of initially ice-rich protoplanets, but significant stripping of gaseous atmospheres appears likely. Hence, it may be difficult to reconcile the observation that many low mass Kepler planets have H/He envelopes with an *in situ* formation scenario that involves giant impacts after dispersal of the gas disc.

3.1. Introduction

The Kepler mission discovered 4723 exoplanet candidates, of which 2302 have been confirmed as bona fide transiting planets (Borucki et al., 2010, 2011; Batalha et al., 2013; Burke et al., 2014; Rowe et al., 2015; Mullally et al., 2015; Coughlin et al., 2016; Thompson et al., 2018). More than 70% of Kepler planets have radii 1 $R_\oplus \leq R_p \leq 4$ $R_\oplus$, such that super-Earths and sub-Neptunes make up a large fraction of the known exoplanet population[1]. A significant number of these planets are found in compact multiplanet systems, such as the 6-planet system Kepler-11 (Lissauer et al., 2011a) and the 5-planet system Kepler-84 (Rowe et al., 2014). The highest multiplicity system detected by Kepler where all planets are confirmed is Kepler-90, with eight planets transiting its host star (Shallue and Vanderburg, 2018). Analyses of the Kepler data to determine occurrence rates of planets shows that systems of Earths and super-Earths with orbital periods < 100 days are common around Solar-type stars (e.g. Fressin et al., 2013; Petigura et al., 2013). A recent analysis suggests that the mean multiplicity of super-Earth systems with periods < 100 days is ~ 3, approximately 1/3 of Sun-like stars host compact planetary systems, and the mean number of planets per star is ~ 1 (Zhu et al., 2018).

The relative numbers of 1- to 8-planet systems discovered via transit detections are dependent on both the intrinsic multiplicities of the systems, and the mutual orbital inclinations of the planets that comprise the systems. In this work, we examine whether or not a simple model of the *in situ*, final stage assembly of planetary systems, involving dynamical instabilities and accretion through giant impacts among a large population of protoplanets after the gaseous protoplanetary disc has dispersed,

[1] All Kepler planetary data used in this chapter are from NASA Exoplanet Archive unless stated otherwise.

is consistent with the Kepler observations.

Previous N-body simulations have considered the formation of compact systems of planets from an earlier stage than we consider here, and include the influence of the protoplanetary disc and subsequent disc-planet interactions (e.g. Terquem and Papaloizou, 2007; Hellary and Nelson, 2012; Coleman and Nelson, 2014; Cossou et al., 2014; Coleman and Nelson, 2016b). One feature of these simulations is that chains of short period planets in mean motion resonances are a common outcome, contrary to what is observed in the Kepler planet population. More recent work, however, has indicated that these resonant chains can become dynamically unstable once the gas disc has dispersed, such that the final stages of planetary assembly involve mutual scattering and collisions between planets (Matsumoto et al., 2012; Izidoro et al., 2017, 2019; Carrera et al., 2019). In addition to breaking the resonant chains, the gravitational scattering also raises the mutual inclinations and eccentricities, and allows under some circumstances for the simulations to produce planetary system multiplicities that are reported to be in agreement with the Kepler data.

In contrast to the migration-driven formation scenario described above, there have also been N-body studies of *in situ* formation in a gas free environment (e.g. Hansen and Murray, 2012; Moriarty and Ballard, 2016; Matsumoto and Kokubo, 2017). Here, the initial conditions consist of numerous protoplanets arranged in an annulus that undergo mutual scatterings and collisions on the way to assembling the final systems. While these calculations are in some ways similar to the final stages of the migration-driven scenarios when the break-up of the resonant chains occurs, they differ in some important respects. For example, the planets do not start in resonance, and the numbers of bodies involved in the collisional evolution is significantly larger. Hence, the number of collisions experienced by a typical planet is also larger during the evolution.

Our approach in this chapter is similar to that used in the aforementioned *in situ* models, except we use a subset of the Kepler 5-planet systems as templates when constructing the initial conditions of the N-body simulations. We reconstruct the surface density distributions of the chosen planetary systems, and use this to define initial conditions consisting of numerous orbiting protoplanets. The approach is therefore similar to the construction of a minimum mass exoplanet nebula model proposed by Chiang and Laughlin (2013). The protoplanet systems are then evolved for 10^7 yr in a gas-free environment. For each of the planetary systems we consider, we perform two sets of simulations. One uses a traditional hit-and-stick accretion prescription when collisions occur, and the other uses a more complex accretion prescription based on hydrodynamical simulations of colliding bodies (Leinhardt

and Stewart, 2012). Hence, we are able to examine the influence of the accretion prescription on the outcomes of the simulations, similar to the recent study by Mustill et al. (2018).

Adopting a more complex collision model also allows us to track the impact energy during collisions, and we use this information to examine possible composition changes that the planets could potentially experience through the removal of volatile components. Using the relations between the collision energy and the final water content of the largest remnant after differentiated bodies composed of rock and water have collided (Stewart and Leinhardt, 2009; Marcus et al., 2010), we determine how much water could be removed from the planets during their collisional evolution. Although some individual collisions would likely lead to significant compositional changes, taken as a whole our results indicate that the compositions of water-rich super-Earths would not change significantly, if their final stages of evolution were similar to those occurring in the simulations. A similar analysis was also used to examine whether or not the impact energies during collisions could potentially remove putative H/He envelopes from the planets, by the conversion of impact energy into heat energy in the cores (Biersteker and Schlichting, 2019), and here we find that very significant erosion of gaseous envelopes should be expected.

The rest of this chapter is structured as follow. In section 3.2 we describe the simulation methods and the set-up of the initial conditions. In section 3.3 we present the main outcomes of our simulations, and in section 3.4 we examine the formation pathways of the co-orbital planets that arise in the simulations. In section 3.5 we post-process the simulation data and examine the changes to compositions that might arise during collisions, and in section 3.6 we examine the stripping of gaseous envelopes that might arise. In section 3.7, we discuss the results from synthetic observation of the final simulated systems, and examine in particular the distribution of system multiplicities and eccentricities that arises. Finally, we discuss our results and draw conclusions in section 3.8.

3.2. *N*-body simulation methods

We use the N-body codes MERCURY (Chambers, 1999) and SYMBA (Duncan et al., 1998) to undertake the simulations presented in this chapter. Both codes use the Mixed Variable Symplectic (MVS) integration scheme (Wisdom and Holman, 1991), but whereas MERCURY handles close encounters by transitioning to a Bulirsch-Stoer method (Press et al., 1992), SYMBA uses the Regularized MVS scheme (Levison

and Duncan, 1994). More importantly for the work presented here, the versions of the two codes we employ handle collisions differently. MERCURY uses a simple hit-and-stick algorithm that conserves the total mass and linear momentum when two bodies collide and accrete into a single object, whereas our version of SYMBA adopts the imperfect accretion algorithm from Leinhardt and Stewart (2012), which we describe below.

3.2.1. Imperfect collision model

For a detailed description of the Leinhardt and Stewart (2012) collision model we refer the reader to that paper (and section 2.1), and here we simply summarise the post-collision outcomes that are generated by it, along with a few salient details about the implementation. We note that the collision model was implemented in SYMBA by the authors. We refer to the more massive body involved in the collision as the target, and the less massive object as the projectile. The range of outcomes includes the following: a perfect merger where a single body is formed with the total mass and momentum of the original two bodies; a single massive body remains whose (largest-remnant) mass is denoted $M_{\rm LR}$, along with collisional debris in the form of low mass 'super-planetesimals' (gravitating particles that are not mutually interacting); two massive bodies remain with masses $M_{\rm LR}$ and $M_{\rm SLR}$ (mass of the second largest remnant), along with collision debris in the form of low mass 'super-particles'; no massive bodies remain and all the mass is in the form of collision debris represented by low mass 'super-planetesimals'. The following notation is used in the description below: $V_{\rm imp}$ is the impact velocity; $V_{\rm esc}$ is the escape velocity from the colliding bodies (or, more accurately, from the combined target mass and interacting mass of the projectile); $Q_{\rm R}$ is the impact energy; $Q^*_{\rm RD}$ is the catastrophic disruption energy, which by definition corresponds to the impact energy when the mass of the largest remnant contains half of the total mass of the colliding bodies; $b_{\rm crit}$ is the critical impact parameter that determines whether or not a collision is grazing ($b \geq b_{\rm crit}$) or non-grazing ($b < b_{\rm crit}$). The collision algorithm consists of a decision tree with the following possible outcomes (section 2.1.2 and figure 2.4):

- When $V_{\rm imp} < V_{\rm esc}$ we have a perfect merger
- When $V_{\rm imp}$ exceeds the threshold for super-catastrophic disruption, both colliding bodies are destroyed and only collisional debris remains
- When $V_{\rm imp}$ exceeds the threshold for catastrophic disruption or erosion, only one massive body remains and the rest of the mass is in the form of collisional

debris. Here, $M_{\rm LR}$, the mass of the largest post-collision remnant, depends on $Q_{\rm R}$ and $Q^*_{\rm RD}$.

- If $V_{\rm imp}$ is smaller than the threshold for erosion and $b < b_{\rm crit}$, then we have partial accretion where the target body gains mass from the projectile, which is itself completely disrupted into a number of low mass 'super-planetesimals'. $M_{\rm LR}$ again depends on $Q_{\rm R}$ and $Q^*_{\rm RD}$

- For $b \geq b_{\rm crit}$, in descending order of the impact velocities, we have the following outcomes that all preserve the mass of the target object and modify the mass of the projectile: hit-and-spray, where the projectile is completely disrupted into debris particles; hit-and-run, where the projectile mass is reduced and the remaining mass goes into debris particles; bouncing collision, where the projectile retains all of its mass and the collision is treated as an inelastic bounce; graze-and-merge collision, where a single body forms containing all the mass of the colliding objects.

The total mass before and after a collision, $M_{\rm Total}$, is conserved, which means the total mass of the post-collision bodies obey the relation

$$M_{\rm T,d} = M_{\rm Total} - (M_{\rm LR} + M_{\rm SLR}), \tag{3.1}$$

where $M_{\rm T,d}$ is the total mass in debris after the collision. The number of debris particles, $N_{\rm d}$ is given by

$$N_{\rm d} = \begin{cases} \max\left(\dfrac{M_{\rm T,d}}{10.0{\rm M}_{\rm Ceres}}, \min\left(38, \dfrac{M_{\rm T,d}}{0.1{\rm M}_{\rm Ceres}}\right)\right), & \text{if } M_{\rm T,d} > 0 \\ 0, & \text{otherwise} \end{cases} \tag{3.2}$$

where ${\rm M}_{\rm Ceres}$ is the mass of Ceres ($\sim 1.5\times10^{-4}{\rm M}_\oplus$). If the values of $M_{\rm T,d}/10.0{\rm M}_{\rm Ceres}$ and $M_{\rm T,d}/0.1{\rm M}_{\rm Ceres}$ are not even integers, they are rounded-up to the nearest even integers. With the known value of $M_{\rm T,d}$ and $N_{\rm d}$, the mass is evenly distributed to each debris particle.

If debris particles are formed after a collision, they are evenly distributed in a circle on the plane of impact at a distance of one Hill radius ($R_{\rm Hill}$) from the collision centre of mass according to

$$\mathbf{r}_{\rm d} = \mathbf{r}_{\rm com} + R_{\rm Hill}\hat{\mathbf{r}}_{\rm d}, \tag{3.3}$$

where $\mathbf{r}_{\rm d}$ is the initial position vector of the debris particles, $\mathbf{r}_{\rm com}$ is the position vector of the collision centre of mass, and $\hat{\mathbf{r}}_{\rm d}$ is the position unit vector for the

evenly distributed debris with respect to the collision centre of mass. The velocities of the debris particles are simply assumed to be 5% larger than $V_{\rm esc}$:

$$\mathbf{V}_{\rm d} = \mathbf{V}_{\rm com} + 1.05 \times V_{\rm esc}\hat{\mathbf{r}}_{\rm d}, \tag{3.4}$$

where $\mathbf{V}_{\rm d}$ and $\mathbf{V}_{\rm com}$ are the initial velocity vector for the debris and the collision centre of mass velocity vector.

The accumulated effect of high energy collisions can lead to the creation of collisional debris in the form of thousands of 'super-planetesimals'. These particles normally get reaccreted by the protoplanets during the simulations, but if, for example, a super-catastrophic collision occurs at the inner edge of our system, then a ring of planetesimals can form which have exceedingly long dynamical life times. This then causes the simulation run times to increase appreciably. To ameliorate this situation, we have introduced a scheme for removing such a ring of particles when it forms. This is motivated by the fact that the collision time in the ring is normally very short, and collisions between the planetesimals would be highly destructive, such that they would be ground down to dust which would then be removed by radiation pressure and/or Poynting-Robertson drag. The scheme calculates the collision time and reduces the masses of the planetesimals on that time scale, until the mass in the ring is negligible and the particles can be removed from the simulation. A more detailed description is given in appendix A.1.

3.2.2. Kepler multi-planet system templates

In this study we have selected a number of Kepler 5-planet multi-planet systems to provide templates for the initial conditions of the simulations, using the following criteria. We are interested in the compact systems, so we have chosen systems in which the known outermost planet has semi-major axis $\leq$ 0.5 au. We have selected those Kepler systems where all five of the known planets are transiting. For example, the Kepler-122 system is not included due to one of its planets, Kepler-122f, being discovered by transit timing variations (TTVs) (Hadden and Lithwick, 2014). When we began this project, Kepler-80 was listed as a 5-planet system (MacDonald et al., 2016), but more recently it has been confirmed as a 6-planet system using deep learning by Shallue and Vanderburg (2018). In spite of this recent announcement, we include this system using the five planets known before 2018. Kepler-296 is a binary system with two stars, Kepler-296A and -296B, that have a projected separation of $\sim$ 70 au (Barclay et al., 2015). All five planets are

orbiting the same star (Kepler-296A). Given that the outermost planet, Kepler-296Af, orbits at ~ 0.255 au, which is only $\sim 0.36\,\%$ of the binary stars separation, the binary should have little influence on the dynamic stability and evolution of the planetary system (Wiegert and Holman, 1997), and hence we include this system.

As described below, we use the Kepler systems to construct individual mass surface density profiles, which are then used to produce initial conditions for the simulations consisting of 20 protoplanets. We impose selection criteria on these initial conditions that include a requirement that the inter-protoplanet separation is not too small or too large (i.e. $5 \leq K \leq 30$), where K is the inter-protoplanet separation measured in units of the mutual Hill radius. This avoids the evolution being dominated by collisions that occur at very early times before dynamical relaxation of the systems has had an opportunity to arise, or the converse where no collisions happen at all. Finally, we require the maximum value of the initial protoplanet mass to be $M_{\rm p} < 6\ {\rm M}_{\oplus}$. After applying these criteria, eight systems were selected to be the templates. As listed in table 3.1, these are Kepler-55, -80, -84, -102, -154, -169, -292 and -296.

3.2.3. Surface density profiles from Kepler systems

In order to construct surface density profiles from the chosen Kepler systems, we need to know the semi-major axes and masses of the planets, which are not provided directly by the observations. The semi-major axis is obtained trivially from Kepler's 3$^{\rm rd}$ law

$$a = \sqrt[3]{\frac{GM_{\star}}{4\pi^2}P^2}. \tag{3.5}$$

where P is the measured planetary period and $M_{\star}$ is the mass of the host star. Numerous suggested relations between the observed planet radius, R_p, and the planet mass, M_p, have appeared in the literature. In this study, we adopt the relation suggested by Lissauer et al. (2011b), based on fitting the Earth and Saturn:

$$M_p = \left(\frac{R_p}{{\rm R}_{\oplus}}\right)^{2.06} {\rm M}_{\oplus}. \tag{3.6}$$

We also considered the relation suggested by Wolfgang et al. (2016) (which is the best-fit relation for the sample of RV-measured transiting sub-Neptunes with $1.5 < R_p < 4\ {\rm R}_{\oplus}$). As discussed later in Section 3.3.1, however, we find that obtaining M_p from R_p using this relation results in some of the selected Kepler systems being themselves dynamically unstable on relatively short time scales, hence we did not

adopt this mass-radius relation in this study.

Once we have M_p and R_p, the internal density, ρ_p, is given by

$$\rho_p = \frac{3M_p}{4\pi {R_p}^3}. \tag{3.7}$$

In order to simplify the collision model for a given simulation, we adjusted the ρ_p values within each individual system to be the same for different planets, obviating us from having to deal with collision outcomes involving planets with significantly different densities. We did this by constructing a mass weighted average of the planetary densities, as follows:

$$\langle \rho_p \rangle = \frac{\sum M_p \rho_p}{\sum M_p}. \tag{3.8}$$

To find the surface density profile for the original Kepler system, we first define an annulus surrounding each planet. Figure 3.1 shows an example planetary system and the annuli associated with each planet, where each annulus is defined by its inner and outer radius. For a general planet i, these are denoted as R_i and R_{i+1}. Here, R_i is taken to be the midpoint between the semi-major axis of the planet, a_i, and its inner adjacent planet, a_{i-1}:

$$R_{i,i\neq 1} = \frac{a_i + a_{i-1}}{2}. \tag{3.9}$$

The innermost boundary is located at

$$R_1 = a_1 - \left(\frac{a_2 - a_1}{2} \right), \tag{3.10}$$

and when the planetary system has n planets, the outermost boundary, R_{n+1}, is at

$$R_{n+1} = a_n + \left(\frac{a_n - a_{n-1}}{2} \right). \tag{3.11}$$

The area of the i-th annulus, A_i is

$$A_i = \pi \left(R_{i+1}^2 - R_i^2 \right), \tag{3.12}$$

and the surface density of the annulus can be calculated using $M_{p,i}$ (the mass of the planet contained in the annulus obtained from equation 3.6) and A_i from equa-

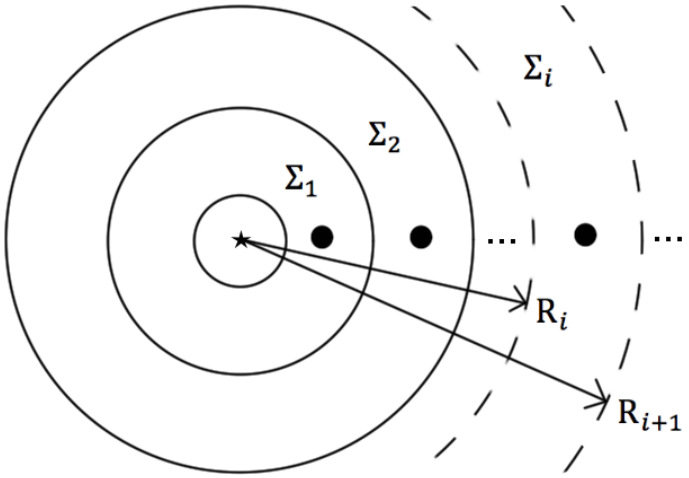

Figure 3.1.: Diagram illustrating method for calculating $\Sigma_{\rm fit}$, the surface density of each of the Kepler templates we have adopted, as described in the text.

tion 3.12:

$$\Sigma_i = \frac{M_{p,i}}{A_i}. \tag{3.13}$$

This gives the surface density at discrete radial locations around each star, and to obtain the surface density as a continuous function we simply fit the five Σ_i values with a smooth function. A 4-th order polynomial can always be found that passes through 5 real data points, but this approach has not been used here because it often gives negative values of Σ at some locations. Instead, we have fitted the Σ_i using four different model functions, namely:

$$\Sigma_{\rm fit}(a) = \begin{cases} c_1 a^{c_2} + c_3 & \text{(3.14a)} \\ c_1 a^3 + c_2 a^2 + c_3 a + c_4 & \text{(3.14b)} \\ c_1 \exp(c_2 a) + c_3 & \text{(3.14c)} \\ c_1 + c_2 \cos(c_4 a) + c_2 \sin(c_4 a), & \text{(3.14d)} \end{cases}$$

where $\Sigma_{\rm fit}(a)$ is the fitted surface density profile as a function of semi-major axis a, equation (3.14a) is the power fitting model; (3.14b) is the polynomial fitting model; (3.14c) is the exponential fitting model; (3.14d) is the Fourier series fitting model. The c_i's are the fitting coefficients. The selection criteria for which model to choose are: 1) the model that provides the best least-squares fit among all models; 2) no negative values between $\Sigma_{\rm fit}(a = R_1)$ and $\Sigma_{\rm fit}(a = R_n)$. Appendix A.2 provides additional details about the coefficients used to obtain $\Sigma_{\rm fit}$ in this study.

3.2.4. Constructing initial conditions for the simulations

The motivation behind this study is to determine whether or not dynamical instabilities and giant impacts among a large population of protoplanets can lead to final planetary systems that are similar to the compact Kepler multi-planet systems. The initial conditions of the N-body simulations consist of 20 protoplanets that are constructed from the $\Sigma_{\rm fit}$ obtained for each of the Kepler systems.

First we assume the semi-major axes of the new protoplanets are distributed following a power law, which is achieved by defining 20 new annuli with appropriate boundaries. The radius of the innermost boundary R_1 and outermost boundary R_{n+1} always remain at the same position as the original Kepler system. The radius of the i-th boundary in between R_1 and R_{n+1} can be calculated by

$$R_{i+1} = f \times R_i, \tag{3.15}$$

where f is a constant distance ratio and is given by

$$f = \sqrt[n]{\frac{R_{n+1}}{R_1}}. \tag{3.16}$$

Once the R_i are obtained from equation 3.15, the semi-major axis of the i-th protoplanet, a_i, is set at the mid-point between R_i and R_{i+1}. The mass, $M_{p,i}$, of the protoplanet in position a_i is calculated according to

$$M_{p,i} = 2\pi R_i \left(R_{i+1} - R_i\right) \Sigma_{\rm fit}(a_i). \tag{3.17}$$

Using this process to find the $M_{p,i}$ values may result in a system with total mass that differs from the original Kepler system, in which case the mass of each protoplanet is scaled appropriately. Figure 3.2 shows the masses of the initial 20 protoplanets with respect to their semi-major axis for each system template. We also assumed the density of the protoplanets are $\rho_p = \langle \rho_p \rangle$ throughout the whole system. With the new value of ρ_p and M_p, the planetary radii, R_p, adopted in the simulations can be obtained from equation 3.7.

The initial eccentricities, e, inclinations, I, arguments of pericentre, ω, longitudes of ascending node, Ω, and mean anomalies, M, also need to be defined when setting the initial conditions of the simulations. The values of e and I are uniformly distributed within a range $0 \leq e \leq e_{\rm max}$ and $0 \leq I \leq I_{\rm max}$, where $e_{\rm max}$ and $I_{\rm max}$ are defined below. The values of ω, Ω, and M are distributed uniformly in the range $0 \leq (\omega, \Omega, M) \leq 2\pi$.

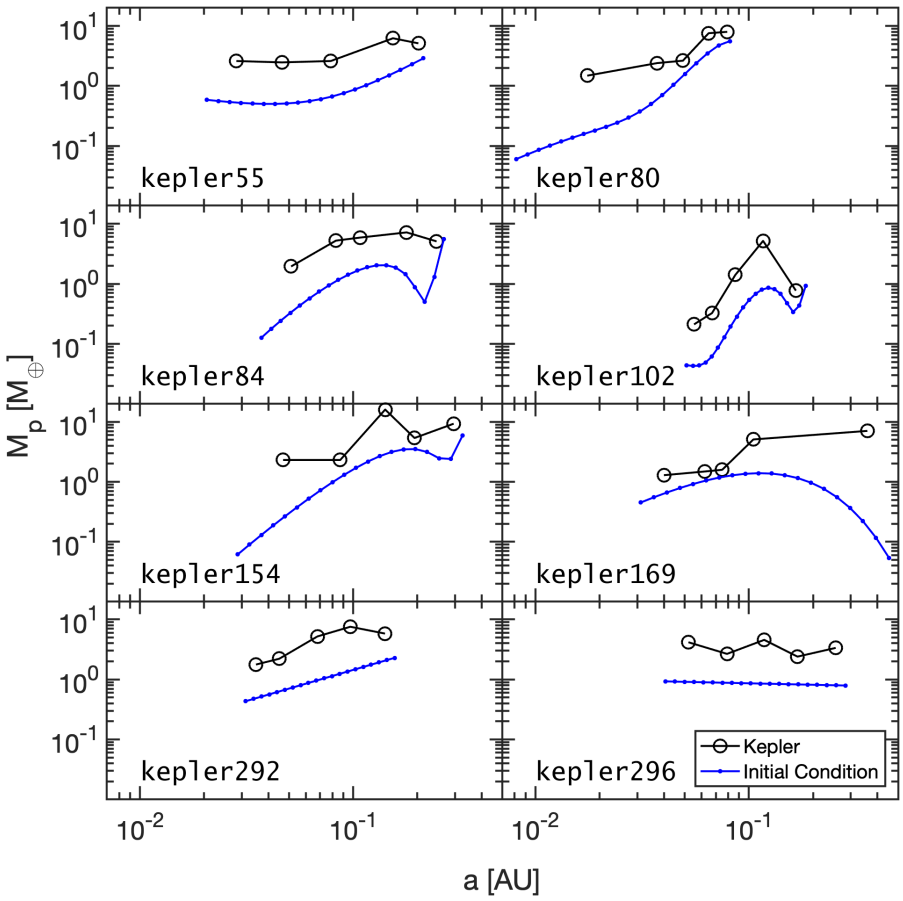

Figure 3.2.: M_p versus a for all 8 Kepler templates. Initial masses of the 20 protoplanets are marked in blue dots, and the original Kepler planet masses are marked in black circles.

Two sets of $e_{\max}$ and $I_{\max}$ values are used here to investigate the effect of the initial eccentricities and inclinations. The first (higher initial value) set has $e_{\max} = 0.02$ and $I_{\max} = 0.01$ rad, while the second (lower initial value) set has $e_{\max} = 0.002$ and $I_{\max} = 0.001$ rad. In each $e_{\max}$-$I_{\max}$ set, 10 simulations were run, with different random number seeds being used to generate the values of ω, Ω, and M. Hereafter, the 'higher set' refers to the runs with ($e_{\max}$, $I_{\max}$)=(0.02, 0.01), and 'lower set' refers to the runs with ($e_{\max}$, $I_{\max}$)=(0.002, 0.001). Each higher and lower set was run using both perfect and imperfect collision models using the MERCURY and SYMBA N-body codes, respectively. Hence, there are 40 simulations for each Kepler system template.

The central bodies of each system have their masses and radii taken from the Kepler data. Each simulation runs for 10^7 yr. The time steps used in the simulations are set to be $1/20^{\text{th}}$ of the shortest orbital period (Duncan et al., 1998).

Table 3.1.: K-values of the selected Kepler 5-planet systems. K_1 to K_4 denote the K-value from the 1st to 4th pair of adjacent planets, respectively, where the 1st pair is the innermost pair and 4th pair is the outermost pair (penultimate and outermost planet). $\bar{K}$ denotes the arithmetic mean of K for the system. Underlined values are the minimum K values in the system, $K_{\min}$. The minimum $K_{\min}$ in the table is 7.2, which is greater than the $K_{\min} = 7.1$ required to give 100% stable rates in a 5-planet system for up to 10^6 years from Wu et al. (2018).

System	K_1	K_2	K_3	K_4	$\bar{K}$
Kepler-55	23.9	25.6	26.7	<u>10.3</u>	21.6
Kepler-80	41.2	14.5	11.4	<u>7.2</u>	18.6
Kepler-84	24.6	<u>11.6</u>	21.0	13.9	17.8
Kepler-102	22.1	19.4	<u>14.8</u>	18.0	18.6
Kepler-154	34.5	17.4	<u>10.9</u>	16.4	19.8
Kepler-169	29.2	<u>12.4</u>	16.8	45.3	26.0
Kepler-292	15.1	20.0	<u>14.4</u>	14.9	16.1
Kepler-296	17.2	16.2	<u>14.8</u>	18.0	16.5

As discussed above in Section 3.2.2, one of the criteria used to constrain our initial conditions is that the mutual separation between neighbouring protoplanets must satisfy $5 \leq K \leq 30$, where K is the inter-protoplanet separation measured in units of the mutual Hill radius. The mutual Hill radius for a pair of adjacent planets is defined by

$$R_{H,i} = \frac{a_i + a_{i+1}}{2} \left(\frac{M_{p,i} + M_{p,i+1}}{3M_\star} \right)^{\frac{1}{3}}. \tag{3.18}$$

The dimensionless number K can then be expressed as

$$K_i = \frac{a_{i+1} - a_i}{R_{Hi}}, \tag{3.19}$$

where K_i is the K-value for the i-th pair of adjacent planets in the system.

For a Kepler planetary system, applying the value of a obtained from equation 3.5, M_p from equation 3.6, and $M_\star$ from the Kepler data, K can be directly calculated by equation 3.19. For the selected Kepler planetary systems in this study (see section 3.2.2), the K_i values for each planet pair and mean K value of each system, $\bar{K}$, are listed in table 3.1. The overall mean K value across all selected systems has the value $\langle \bar{K} \rangle \approx 19.4$. This value is about the same as the typical average K value for Kepler multiplanetary systems (see Section 3.1)

Finally, we comment that the initial conditions of the simulations presented in

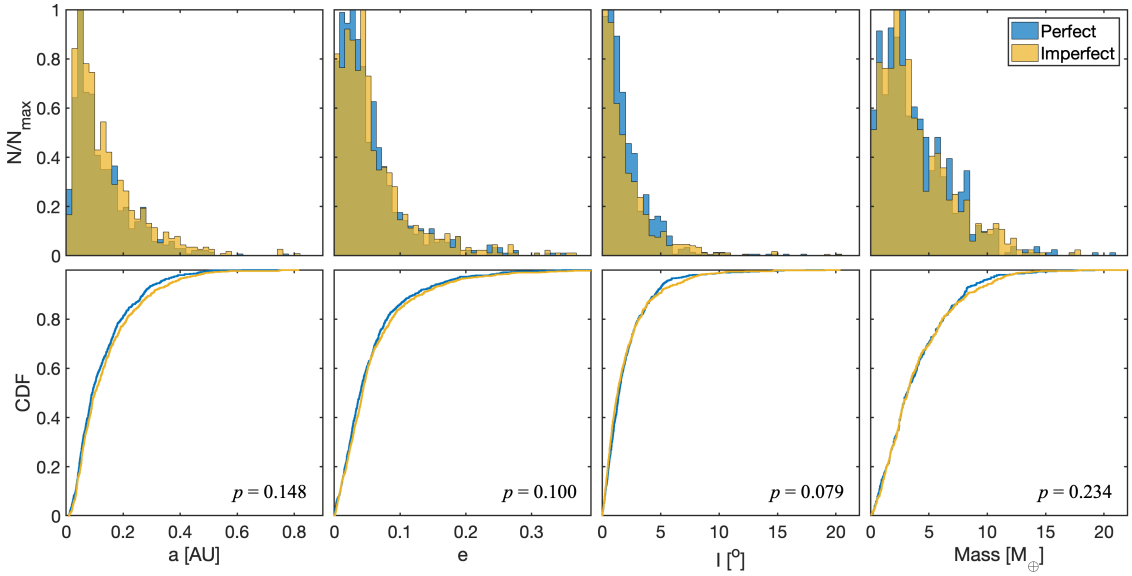

Figure 3.3.: The normalised distributions (top panel) and the cumulative distribution functions (CDFs, bottom panel) of the semi-major axes, eccentricities, inclinations, and planet masses of the planets obtained in all simulations. Perfect collision simulations are shown in blue, and imperfect collision simulations are shown in yellow.

this chapter represent the state of the system after substantial evolution has already taken place, and once the gaseous protoplanetary disc has been dispersed. For discussion of possible scenarios leading to these initial conditions, involving the accretion of planetesimals, boulders and/or pebbles onto planetary embryos embedded within the gas disc, we refer the reader to the following papers that present the results of N-body simulations of these earlier epochs of planet formation (Coleman and Nelson, 2014, 2016b; Lambrechts et al., 2019; Izidoro et al., 2019).

3.3. Results

3.3.1. Stability of the original Kepler multi-planet systems

Before embarking on a study of the formation of the Kepler systems considered in this chapter, we begin by considering the dynamical stability of the observed systems themselves. This acts as a consistency check on the mass-radius relation used to construct the initial conditions for the formation simulations, given by equation (3.6) (Lissauer et al., 2011b). Since the first multiplanet systems discovered by Kepler were confirmed, a number of mass-radius relations have been suggested. For example, Wu and Lithwick (2013) suggested $M_p = 3\mathrm{M}_\oplus(R_p/\mathrm{R}_\oplus)$, Weiss and Marcy

(2014) suggested $M_p = 2.69\mathrm{M}_\oplus(R_p/\mathrm{R}_\oplus)^{0.93}$, and Wolfgang et al. (2016) suggested $M_p = 2.7\mathrm{M}_\oplus(R_p/\mathrm{R}_\oplus)^{1.3}$. Clearly, given that the masses obtained from each of these relations vary for specific values of the planetary radii, stability of the confirmed system is not guaranteed to hold under all these relations. Checking the stability hence provides some constraint on the mass-radius relation that applies.

We carried out a stability check for all Kepler 5-planet systems, including those which do not obey the selection criteria mentioned in Section 3.2.2, comparing the relation suggested by Wolfgang et al. (2016) and that provided by Lissauer et al. (2011b). We performed N-body simulations using MERCURY, and adopted initial conditions that assumed the system planets are initially on circular and coplanar orbits ($e = 0$ and $I = 0$ for all planets). The initial values of the mean anomalies of the planets were assigned randomly, and 10 different realisations were run for each system. We found that the relation from Lissauer et al. (2011b) provides stable systems for all 5-planet systems over a 10 Myr run time, while the mass-radius relation provided by Wolfgang et al. (2016) fails to produce stable systems for some Kepler systems over the same time scale (e.g. the systems Kepler-32 and -33, which were validated by Lissauer et al. (2011b)). For this reason, the relation provided by Lissauer et al. (2011b) is the one we used to construct the initial conditions for the 20-protoplanet simulations described in the following sections.

3.3.2. Results of the formation simulations

To recap, two sets of N-body simulations were performed for each Kepler template, one assuming perfect accretion using a simple hit-and-stick model, and the other adopting the imperfect accretion algorithm of Leinhardt and Stewart (2012). For each Kepler template, we considered two initial distributions of the eccentricites and inclinations, a 'High set' and a 'Low set', and for each of these we computed 10 different realisations of the initial conditions by varying the random number seeds used to create the initial conditions. Hence, we ran 40 N-body simulations for 10 Myr for each Kepler template.

We begin our discussion of the results by first considering how the simulation outcomes considered as a whole vary when considering the perfect and imperfect accretion prescriptions. We then look at the simulations in more detail by considering how the outcomes vary between the different Kepler templates, focussing on the resulting planet masses, orbital elements, period ratios, K-values and system architectures that emerge from the simulations.

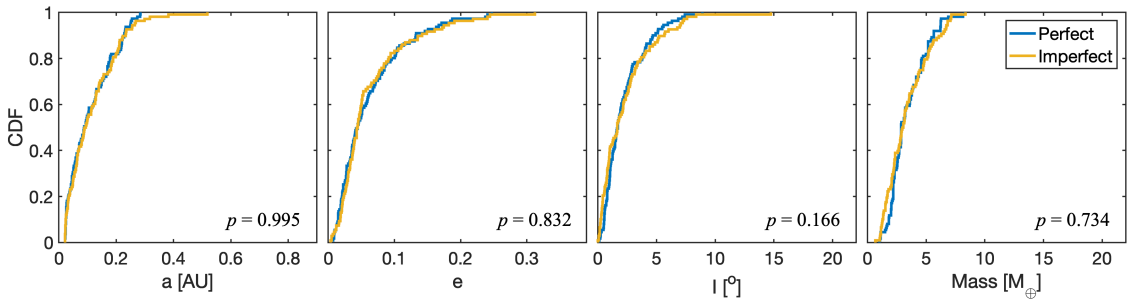

Figure 3.4.: Cumulative distributions, from all simulations of the `Kepler55` template, of the same 4 parameters listed in figure 3.3. They are (in order from left to right) a, e, I, and M_p. The blue and yellow lines are for the perfect collision and imperfect collision simulations, respectively.

Comparison between perfect and imperfect accretion across all runs

The distributions of the semi-major axes, eccentricities, inclinations and masses are shown in the histograms and cumulative distribution functions (CDFs) in figure 3.3. Later in the chapter we discuss the mutual separations between pairs of planets, and the K-values are shown in figure 3.11. By-eye inspection suggests that the distributions are in good agreement when comparing the perfect and imperfect accretion models, and applying the Kolmogorov-Smirnov test (K-S test) yields p-values of 0.148 for the semi-major axis distributions, 0.100 for the eccentricities, 0.079 for the inclinations, and 0.234 for the masses. Hence, the null hypothesis that the data plotted in figure 3.3 for the perfect and imperfect collision simulations are drawn from the same underlying distribution cannot be rejected with a $> 95\%$ confidence level.

Agreement between the perfect and imperfect accretion runs can also be seen when looking at individual system templates. For example, figure 3.4 shows the CDFs for the same parameters shown in figures 3.3, but only for the `Kepler55` system template, and again it can be seen that the distributions are very similar. Here the K-S test yields p-values of 0.995 for the semi-major axis distributions, 0.832 for the eccentricities, 0.166 for the inclinations, and 0.734 for the masses.

In summary, based on the global properties of the final planetary systems that are formed, we can conclude that the differences produced by the perfect and imperfect accretion prescriptions are small, and do not have a statistically significant influence on the outcomes of the simulations.

Instabilities and multiplicities

All simulations resulted in dynamical instabilities that led to mutual scattering and giant impacts. Figure 3.5 shows the distributions of the multiplicities of all final planetary systems. The maximum number of planets remaining after 10 Myr was 12 and the minimum was 3. No single or double planet systems were formed. Having 4 or 5 planets remain in the system is the most common outcome. Our multiplicity distribution appears to agree with the distribution obtained by Hansen and Murray (2013), where they also obtain a minimum multiplicity of 3, and a peak in the multiplicity distribution at 4 or 5 planets. The mean value of the number of planets obtained in the imperfect collision simulation is $\langle N_I \rangle = 5.21$ and for perfect collisions it is $\langle N_P \rangle = 5.06$. For the higher initial value set $\langle N_{\rm H} \rangle = 5.05$, and for the lower initial value set $\langle N_{\rm L} \rangle = 5.22$. It is noteworthy how close these values are to 5, given that our template systems all contain 5 planets, indicating that the initial conditions constructed from the templates are able to reproduce the desired multiplicity on average. In a recent analysis of the Kepler data, Zhu et al. (2018) concluded that the mean number of super-Earths in compact systems around solar-type stars is approximately 3, with the fraction of stars hosting planetary systems being approximately 0.3. This suggests that the 5-planet systems we have chosen for this study may not be representative of the Kepler planets as a whole, even if we allow for the fact that the Kepler systems contain unseen planets by virtue of them being on orbits that are inclined to the line of sight.

The K-S test applied to the CDFs derived from the data in figure 3.5 gives a p-value of 0.997 for the perfect and imperfect collision models, and 0.999 for the runs with the higher and lower initial eccentricity/inclination values. The small difference between $\langle N_I \rangle$ and $\langle N_P \rangle$, together with the large p-value, shows that our runs are in accord with the conclusions reached by Mustill et al. (2018), namely that assuming either perfect or imperfect collisions has little impact on the final multiplicities. The similar values for $N_{\rm H}$ and $N_{\rm L}$ shows that the initial value of e and I also has a limited impact on the final multiplicities, at least for the range of values adopted here.

Figure 3.6 shows the CDF for the occurrence times of all giant impact events detected during the imperfect collision simulations. More than 90% of the giant impacts happened before 1 Myr (the white area in the figure), and 50% of the impacts occurred within 10^4 yr. Given that the planetary systems are centred around $a \sim 0.1$ au, this latter figure corresponds to $\sim 3 \times 10^5$ dynamical times, indicating that the initial conditions do not result in excessively short accretion times. Instead,

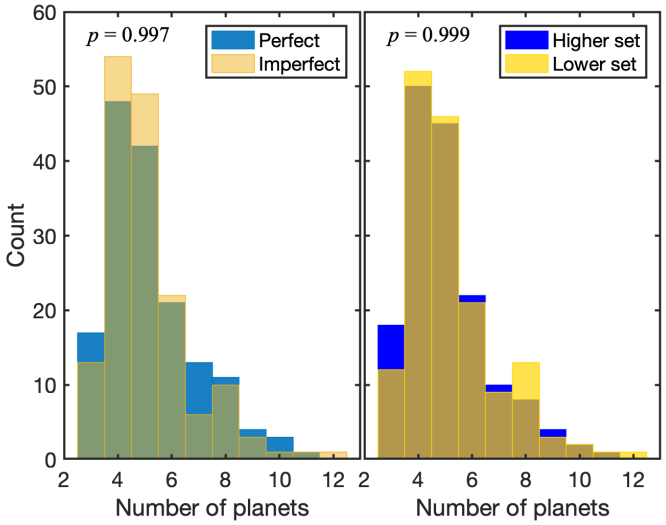

Figure 3.5.: Multiplicity distributions from all simulations. The left panel compares the distributions obtained in the perfect and imperfect collision simulations. The right panel compares the distributions obtained in the high and low initial eccentricity/inclination simulations (see section 3.2.4 for the definitions of these simulation sets).

the systems have time to undergo substantial dynamical relaxation during the epoch of accretion. Furthermore, the fact that only 10% of the collisions occur after 1 Myr indicates that our run times of 10 Myr are long enough to have formed long-term stable systems in most cases. However, the fact that some collisions are occurring at late times also indicates that some of our final planetary systems would have evolved further if the integrations had been extended. Finally, we note that with 90% of the collisions occurring in the simulations within 1 Myr, this implies that if the protoplanets we consider in the initial conditions were formed within the life time of the gaseous protoplanetary discs, then substantial collisional evolution would have likely occurred while the gas was still present as disc life times are typically 3 Myr (Haisch et al., 2001). Ogihara et al. (2015) have shown that under such conditions the effects of migration cannot be ignored and strongly influence the architectures of the resulting systems.

Eccentricities, inclinations and masses

Table 3.2 lists the mean values of e, I and the K-values for each subset of runs associated with each of the Kepler templates. Averaging has been performed over the final outcomes of the 10 simulations associated with perfect/imperfect accretion,

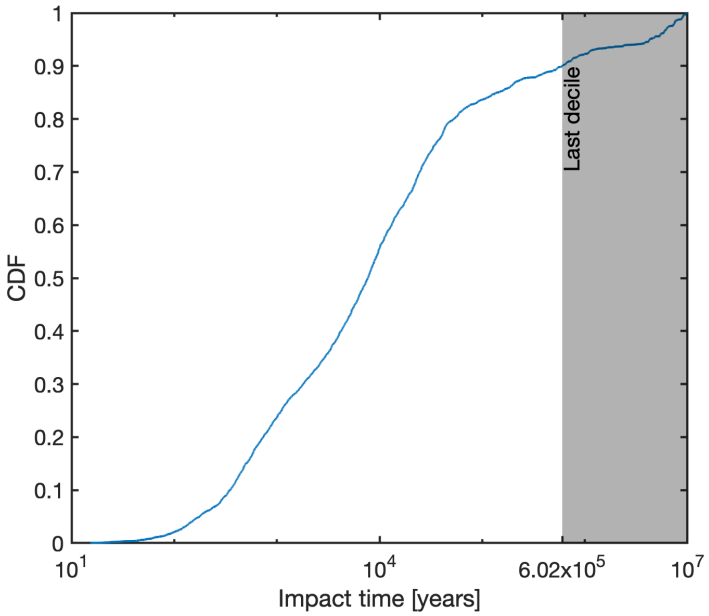

Figure 3.6.: Cumulative distribution of all the giant impact events with respect to time during the imperfect collision simulations. Grey area denotes the latest 10% of the collisions.

and high and low initial eccentricities/inclinations. While there is some variation of the mean eccentricities and inclinations when comparing the different initial eccentricities/inclinations and the accretion prescription (particularly for `Kepler80`), the largest variation is observed when comparing the different Kepler templates. For example, the `Kepler55` runs all give $\langle e \rangle \sim 0.06$, where the `Kepler296` runs give higher values distributed around $\langle e \rangle \sim 0.09$.

The top and middle panels of figure 3.7 show the values of e and I with respect to semi-major axis a from four of the simulation templates performed with the imperfect collision model (see figure A.4 in appendix A.3 for the results of the other 4 templates that adopt imperfect collisions, and figures A.6 and A.8 for all 8 templates that use the perfect collision model). The plots show that the final distributions arising from the higher set (blue triangles) and the lower set (yellow diamonds) are very similar. This is not surprising as the initial eccentricities and inclinations in all of these runs are considerably smaller than the mean values at the ends of the simulations. Hence, the final values are determined by planet-planet scattering and collisional damping, and little memory is retained of the original eccentricity and inclination values. This would not be the case if the initial eccentricities and inclinations had been comparable to or larger than the values obtained from dynamical relaxation (Matsumoto and Kokubo, 2017). It is common to see

Table 3.2.: Mean values of e, I and K-values from the different simulation subsets from all 8 Kepler templates. The numbers in parentheses are the standard deviations about the respective means.

Kepler System	Imperfect Higher	Imperfect Lower	Perfect Higher	Perfect Lower	All Higher	All Lower	All Imperfect	All Perfect	All Sets
				$\langle e \rangle$ $(\times 10^{-2})$					
55	$6.32_{(5.71)}$	$6.27_{(5.50)}$	$6.03_{(4.46)}$	$6.34_{(5.97)}$	6.18	6.31	6.30	6.19	6.24
80	$3.21_{(4.47)}$	$5.09_{(5.50)}$	$1.85_{(2.25)}$	$1.88_{(1.90)}$	2.53	3.49	4.15	1.87	3.01
84	$7.02_{(4.86)}$	$5.20_{(4.56)}$	$6.32_{(6.14)}$	$5.91_{(4.39)}$	6.67	5.56	6.11	6.12	6.11
102	$5.55_{(6.64)}$	$3.74_{(2.81)}$	$5.64_{(4.46)}$	$6.27_{(7.30)}$	5.60	5.01	4.65	5.96	5.30
154	$8.46_{(6.65)}$	$5.25_{(5.68)}$	$7.23_{(7.02)}$	$7.12_{(7.17)}$	7.85	6.19	6.86	7.18	7.02
169	$4.42_{(2.72)}$	$5.49_{(6.22)}$	$4.68_{(4.06)}$	$4.70_{(6.92)}$	4.55	5.10	4.96	4.69	4.82
292	$5.43_{(3.08)}$	$5.04_{(4.00)}$	$6.65_{(5.20)}$	$5.72_{(3.60)}$	6.04	5.38	5.24	6.19	5.71
296	$10.28_{(9.86)}$	$11.40_{(11.05)}$	$7.60_{(5.72)}$	$9.90_{(6.64)}$	8.94	10.65	10.84	8.75	9.80
				$\langle I \rangle$ $(\times 10^{-2})$ [rad]					
55	$4.34_{(4.85)}$	$3.74_{(2.90)}$	$4.31_{(3.55)}$	$3.47_{(2.61)}$	4.33	3.61	4.04	3.89	3.97
80	$1.18_{(1.20)}$	$1.67_{(2.45)}$	$2.72_{(3.89)}$	$1.41_{(1.96)}$	1.95	1.54	1.43	2.06	1.74
84	$3.38_{(2.86)}$	$3.63_{(3.51)}$	$3.18_{(3.61)}$	$3.97_{(4.39)}$	3.28	3.80	3.51	3.58	3.54
102	$2.52_{(3.03)}$	$2.38_{(2.45)}$	$3.65_{(4.30)}$	$3.93_{(3.64)}$	3.08	3.15	2.45	3.79	3.12
154	$5.00_{(5.68)}$	$5.62_{(3.82)}$	$4.94_{(6.94)}$	$5.38_{(6.33)}$	4.97	5.50	5.31	5.16	5.23
169	$2.38_{(1.92)}$	$3.48_{(3.41)}$	$3.42_{(2.32)}$	$3.00_{(2.30)}$	2.90	3.24	2.93	3.21	3.07
292	$4.19_{(2.95)}$	$2.94_{(2.66)}$	$2.33_{(1.87)}$	$2.87_{(2.42)}$	3.26	2.90	3.56	2.60	3.08
296	$5.90_{(3.56)}$	$7.89_{(7.91)}$	$5.32_{(2.76)}$	$5.90_{(4.67)}$	5.61	6.90	6.89	5.61	6.25
				$\langle K \rangle$ $[R_{\rm H}]$					
55	$21.1_{(6.43)}$	$21.2_{(6.91)}$	$22.6_{(6.25)}$	$20.9_{(7.17)}$	21.8	21.1	21.2	21.7	21.4
80	$21.5_{(11.37)}$	$22.8_{(11.09)}$	$20.6_{(10.40)}$	$19.0_{(10.40)}$	21.0	20.9	22.2	19.8	21.0
84	$21.2_{(4.80)}$	$19.2_{(5.42)}$	$20.8_{(4.79)}$	$19.8_{(4.81)}$	21.0	19.5	20.2	20.3	20.3
102	$21.4_{(6.54)}$	$19.1_{(6.01)}$	$20.6_{(5.97)}$	$21.9_{(4.59)}$	21.0	20.4	20.3	21.1	20.7
154	$20.7_{(6.82)}$	$19.0_{(6.46)}$	$20.4_{(7.43)}$	$18.7_{(8.16)}$	20.6	18.9	19.9	19.6	19.7
169	$21.1_{(6.23)}$	$22.5_{(7.61)}$	$22.2_{(6.74)}$	$21.9_{(9.70)}$	21.6	22.2	21.8	22.0	21.9
292	$19.5_{(3.87)}$	$18.5_{(4.23)}$	$18.7_{(4.40)}$	$17.3_{(5.28)}$	19.1	17.9	19.0	18.0	18.5
296	$24.3_{(7.47)}$	$22.9_{(9.15)}$	$22.4_{(5.69)}$	$21.4_{(4.37)}$	23.4	22.2	23.6	21.9	22.8

that e and I have relatively high values near the inner and outer edges of the systems, perhaps best illustrated by the `Kepler102` and `Kepler169` templates. This feature was already noted by Hansen and Murray (2013) in their study of *in situ* formation of super-Earths, and arises because bodies at the edge of the initial annulus of protoplanets are scattered outwards and do not experience collisions that tend to damp the eccentricities and inclinations.

The bottom panels of Figure 3.7 show the final planet masses versus their semi-major axes. A striking feature of these plots is how the simulated systems (denoted by blue triangles for higher set runs, and yellow diamonds for lower set runs) generally match the observed Kepler systems (denoted by black circles joined by solid

lines). Hence, based on this comparison alone, it is reasonable to conclude that the initial conditions and formation histories that we simulate here might be reasonable approximations to those that applied to the actual Kepler systems we have used as templates. The exception is `Kepler169`, where the outer regions of these simulations failed to generate significant collisional growth because the initial masses of the protoplanets there, generated by the method mentioned in section 3.2.4, were too small (giving initial K-values > 15), leading to instability times longer than the 10 Myr simulation run times.

In view of this, we extended the run times of the `Kepler169` template simulations to 100 Myr, using the perfect accretion routine (see figure A.10 in appendix A.3 for the comparison). As expected, the outer regions of the systems experienced increased growth and provided better agreement with the original Kepler masses. This suggests that a better strategy for future work would be to run simulations for a set number of orbits measured at the outer edges of the systems, instead of a fixed number of years as was done in this work.

Although the final planet masses in the simulations match their Kepler templates on average, it is worth noting that when we consider the CDF of planet masses later in this chapter, and compare it with that obtained from the original Kepler template systems (see figure 3.21), the agreement is not good because the simulations produce a range of planetary systems, some of which have higher multiplicity than 5 and hence contain planets with relatively low masses.

As with the eccentricities and inclination distributions discussed above, there are no systematic differences in the final planetary masses when comparing the high and low initial eccentricity/inclination subsets of runs.

The high values of e and I at the edges of the system due to planets being scattered but experiencing fewer collisions there, discussed above and noted by Hansen and Murray (2013), can also be seen in Figure 3.8, which shows the final planets from all runs in the a-e plane (top panel), a-I plane (middle panel), and a-$M_{\rm p}$ plane (bottom panel). We also see from the lower panel that higher mass planets occupy the centre of the a-$M_{\rm p}$ plane, where collisional growth occurs more frequently, with lower mass planets being present at the edges of the annuli where collisions occur less frequently. Figure 3.9 shows that e and I are strongly correlated, as expected for systems that have undergone dynamical relaxation.

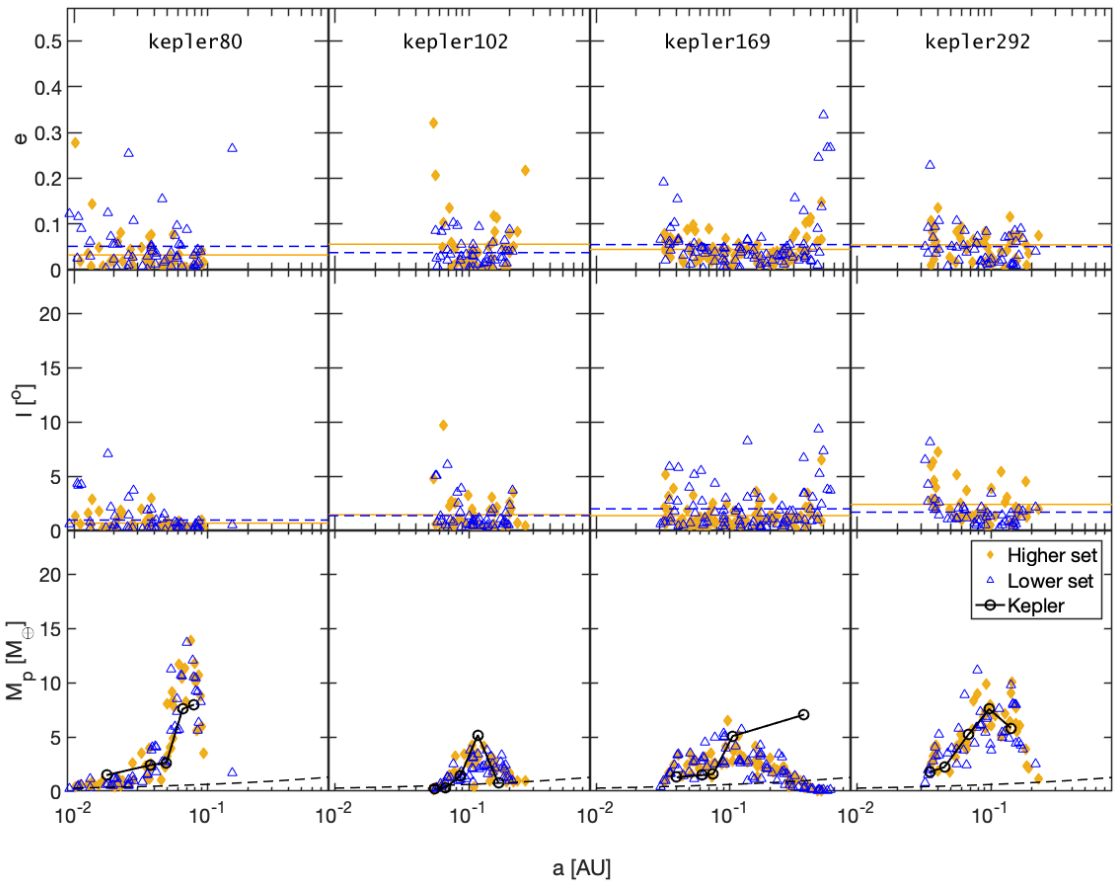

Figure 3.7.: All imperfect collision simulation results from (from left to right) `Kepler80`, `102`, `169`, and `292` templates. The scatter plots show the eccentricities (top panel), inclinations (middle panel), and planet masses (bottom panel) with respect to their semi-major axis. The orange-diamonds data are from the higher initial eccentricity set, and the blue-triangle data are from the lower initial eccentricity set. The horizontal lines in each subplot show the mean values of the data in their respective colour (also plotted as solid lines for the higher initial eccentricity set and dashed lines for the lower initial eccentricity set). The black circles in the bottom panel denote the masses and semi-major axes of the observed Kepler planets. The black dashed lines indicate the detection limit applied when undertaking the synthetic transit observations described in section 3.7.1.

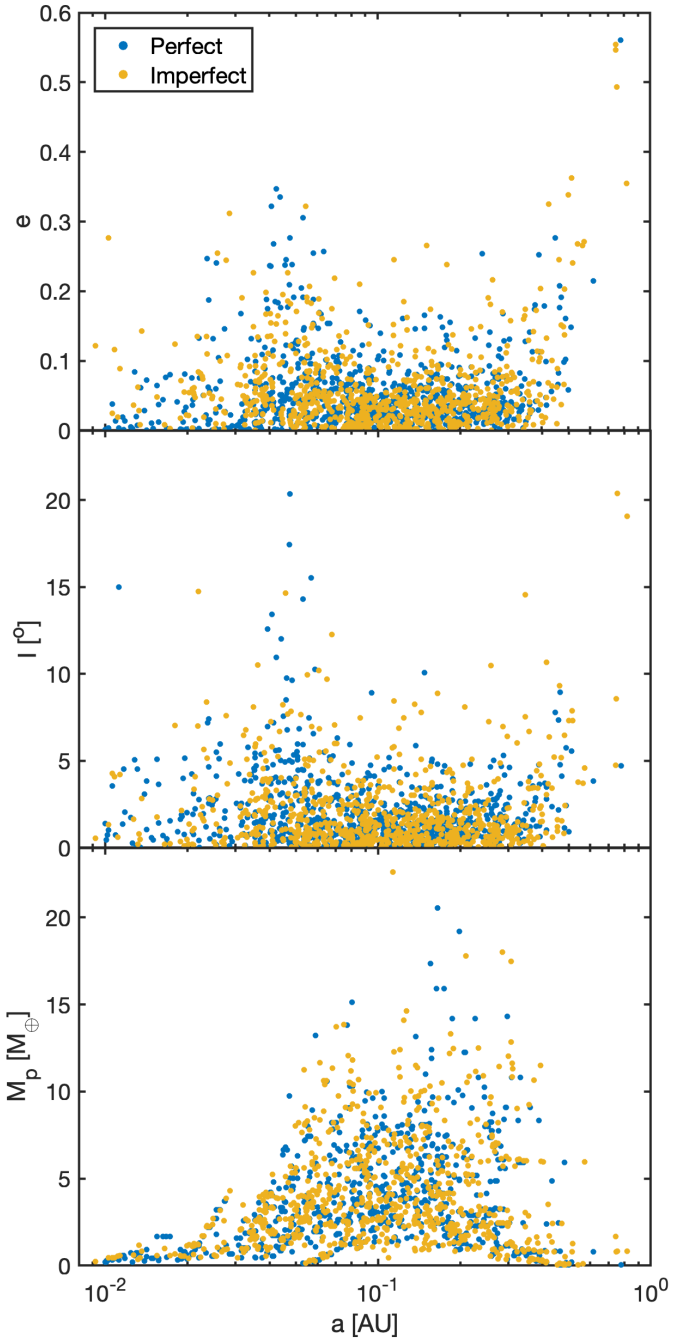

Figure 3.8.: Scatter plots comparing e, I and $M_{\rm p}$ as a function of a arising from perfect (blue points) and imperfect (yellow points) collision simulations.

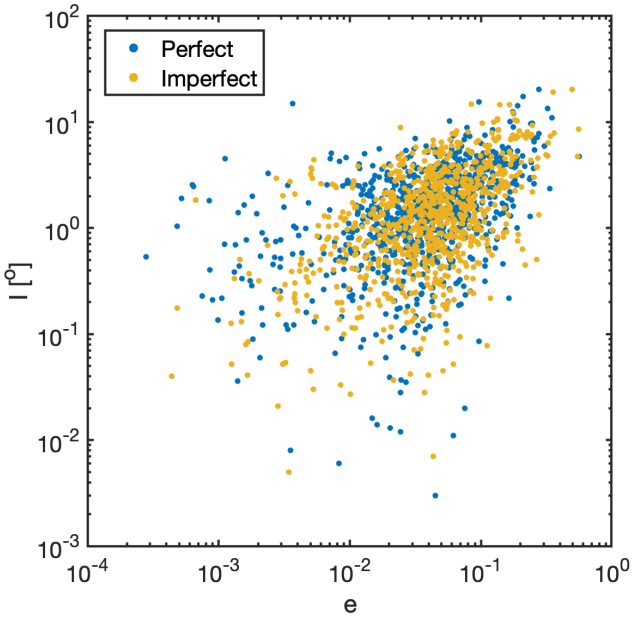

Figure 3.9.: Scatter plot showing correlation between e and I from all runs. Perfect collision results are shown by blue points, and imperfect collision results are shown by yellow points.

Period ratios and K-values

Figure 3.10 shows that the perfect and imperfect accretion simulations provide similar cumulative distributions of the period ratios between neighbouring planet pairs. We can compare these with the distribution of period ratios for the actual Kepler multi-planet systems. In section 3.7.2 below, we also compare the period ratios obtained from the simulations when they are synthetically observed with the Kepler data, but here we focus on the intrinsic period ratios. For period ratios smaller than 4:3, we see that the Kepler data shows an excess compared to the simulations. A K-S test performed on a subset of period ratios between 5:4 and 4:3 gives p-values of 1.69×10^{-4} and 0.029 when comparing the Kepler data with the perfect and imperfect collision simulations, respectively, demonstrating that the distributions are different. Hence, some process occurred during the formation of at least some Kepler systems that allowed the survival of more compact architectures, which are nonetheless non-resonant. Dynamical relaxation and collisional evolution in the absence of any dissipative process clearly results in such closely separated planet pairs being destabilised, suggesting those Kepler systems with particularly compact configurations formed in a dissipative environment and did not undergo dynamical instability in spite of the close proximities of the planets. One such system not considered here

that displays this property is Kepler-11 (Mahajan and Wu, 2014).

The full Kepler data set, without any limits in period ratio being applied, clearly contains too many planet pairs with large period ratios compared to the simulation intrinsic outcomes, and a K-S test comparing the data and simulations results in p-values < 0.05. Again, such system architectures do not naturally arise from a formation scenario in which even a wide annulus of protoplanets undergoes dynamical instability and collisional growth, since this mode of evolution results in neighbouring planets being separated by ~ 20 mutual Hill radii (see discussion below). Instead, additional processes would need to be invoked which either cause the initial distribution of protoplanets to have a more complex structure involving concentrations around particular orbital radii, or which involve orbital migration because formation occurred in the presence of either a gas or planetesimal disc. Considering planet pairs in the Kepler data with a maximum period ratio of 3:1 results in much better agreement between the observational data and simulations, with p-values of 0.078 and 0.120 for perfect and imperfect collisions respectively. However, we also note here that the synthetically observed systems, described in section 3.7.2, provide a distribution of period ratios that is quite different to that obtained from the simulations directly, due to the fact that mutual inclinations between the planets lead to some planets not being detected during the observations. This has the effect of increasing the numbers of systems with large period ratios.

One feature within the Kepler multiplanet systems that the simulations do not reproduce particularly well is the known peaks in occurrence rates of planet pairs just outside of the 3:2 and 2:1 resonances (seen in particular as a flattening and then rise in the CDF at around period ratio 2:1 in figure 3.10). The cumulative distribution for the simulation data shows a very modest inflection around the 2:1 resonance, but it is not as pronounced as in the Kepler data, and is not statistically significant. It is noteworthy that Petrovich et al. (2013) were able to reproduce the resonance features using 3-body integrations that resulted in final systems of two planets for planet masses $20 \leq M_{\rm p} \leq 100$ $M_\oplus$. The end states of our simulations always have more than two planets, and the final planet masses are typically < 10 $M_\oplus$. These properties likely serve to reduce the prominence of features in the period ratio distribution near first order mean motion resonances, and leave open the question of what dynamical processes have given rise to the near-resonance features in the period ratio distributions of the Kepler planets.

The distributions of the K-values from all our simulations can be seen in figure 3.11. Both collision models result in similar distributions, with a p-value of 0.54, with 50% of systems having $10 \leq K \leq 20$, and the maximum value of K being

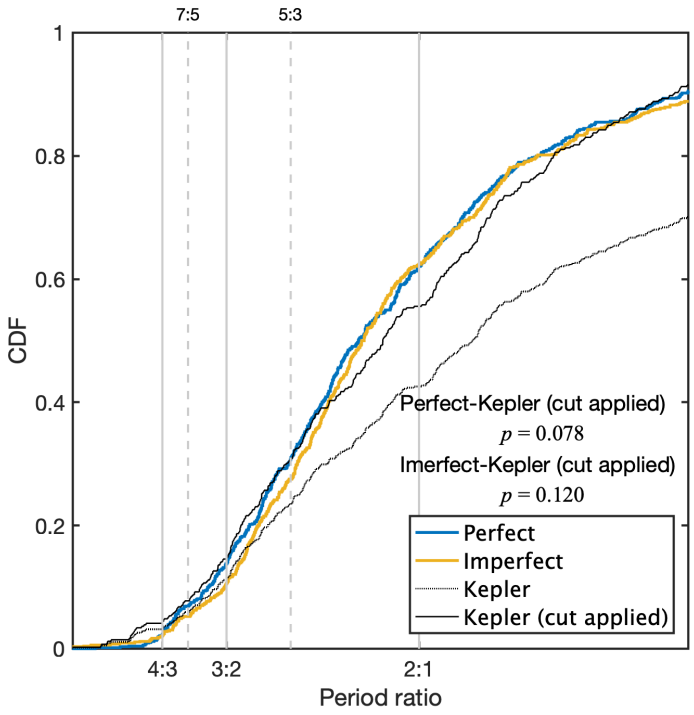

Figure 3.10.: CDFs of the period ratios of all adjacent planet pairs. The blue and yellow lines correspond to the perfect and imperfect collision models, respectively. The black dotted line includes all the original Kepler planet pairs. The solid black line shows the original Kepler planet pairs with a cut off for period ratio > 3.

$\sim$ 50. Although not obvious in the plot, the number of planets with $K < 10$ decreases to zero and then rises again close to $K = 0$, with these latter planets surviving because they are protected by a 1:1 resonance. Wu et al. (2018) have presented similar simulation results on this feature in the K-value distribution and its relation to the 1:1 resonance (period ratio < 1.05), and we discuss these co-orbital systems further in the next section. Figure 3.12 shows the value of K for each planet pair using the same four Kepler templates shown in figure 3.7. The distributions of the K-values are similar, independent of whether we consider the high or low eccentricity/inclination set. And they also show similar K-values compared to the original Kepler systems (shown by the black circles joined by lines), based on the adopted mass-radius relation, although again the `Kepler169` system is an exception (see figure A.11 in appendix A.3 for a comparison to systems that were evolved for 100 Myr instead of 10 Myr).

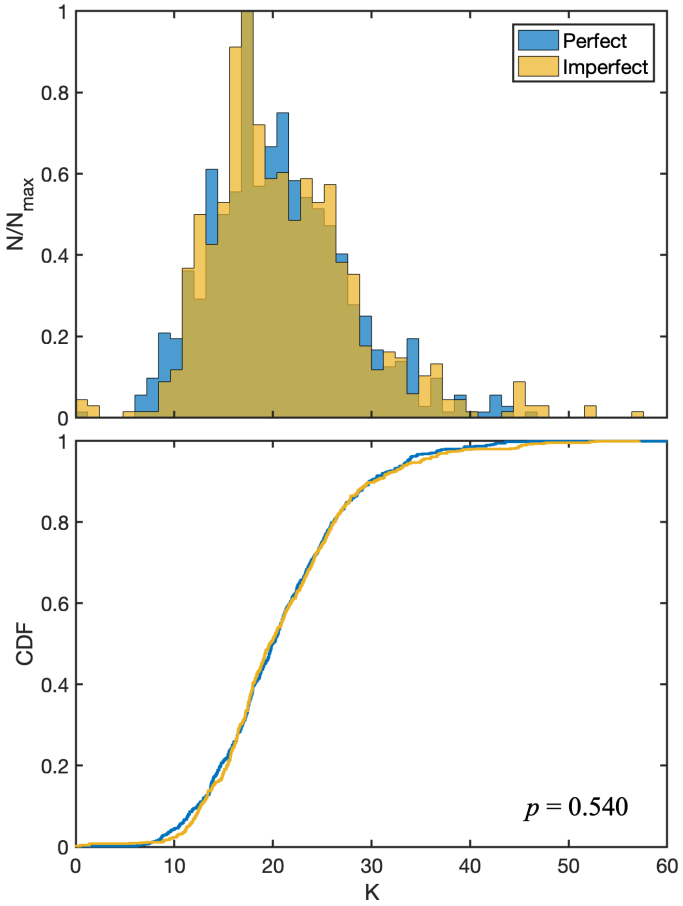

Figure 3.11.: The normalised distributions (top panel) and the cumulative distributions (bottom panel) of the K-value of the planet pairs obtained in all our simulations. Perfect collision simulations are shown in blue, and imperfect collision simulations are shown in yellow.

3.4. Co-orbital planet pairs

In section 3.3.2, we stated that a small number of planets have very small K-values, and these are shown in figure 3.11. Further investigation has shown that these planets have been captured into 1:1 co-orbital resonances, and these co-orbital planets make up about 1% (4 out of 320 simulations) of the total number of planet pairs.

3.4.1. Stability

In spite of the very small K-values, the 1:1 resonance protects co-orbital planet pairs from instability. In general both stable tadpole orbits, which involve libration around the L4/L5 points, and horseshoe orbits are permitted (Dermott and Murray,

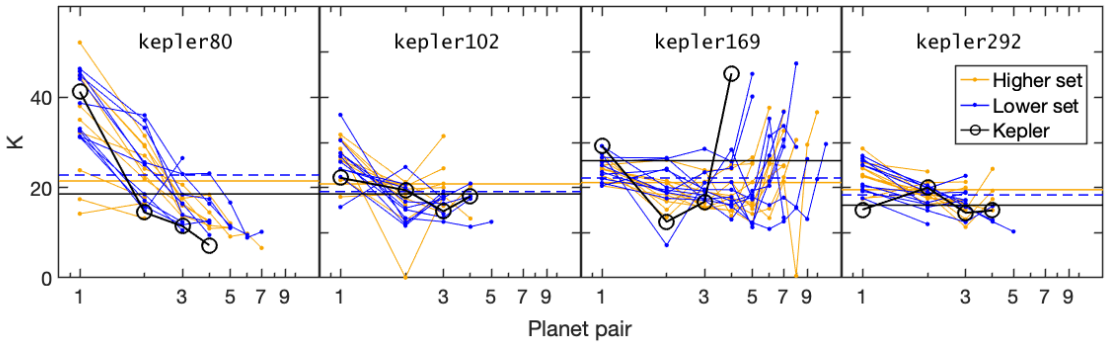

Figure 3.12.: All K-values for neighbouring planet pairs from imperfect collision simulations of (from left to right) `Kepler80`, `102`, `169`, and `292` templates. The orange diamonds are from the higher initial eccentricity set and the blue triangles are from the lower initial eccentricity set. The horizontal lines in each subplot are the mean values of the data in their respective colour (also plotted as solid lines for the higher initial eccentricity set and dashed lines for the lower initial eccentricity set). The black circles denote the K-values of the original planets pairs, as listed in table 3.1.

1981a,b), and we see examples of both of these orbit types in the simulations. As the simulation run times are 10 Myr, and the co-orbital pairs are found to form early in some simulations, we find tadpole and horseshoe orbits which are stable for 9.5×10^6 years. This is in spite of the co-orbital pairs being in systems of high multiplicity, where the resonance configuration is subject to external perturbations. Previous studies have shown that co-orbital planet pairs can be stable for up to 10^9 yr (Tabachnik and Evans, 2000). Figure 3.13 shows an example of the semi-major axis versus time during the last 10 years of one simulation (the left panel) and the corresponding orbital trajectory of the 1:1 resonance planet pair from the same model in a frame that co-rotates with one of the planets (the right panel). From the semi-major axis plot, we can see the 2 planets undergo a periodic exchange of radial location, and the co-rotating plot shows that the orbit is a tadpole orbit.

In other simulations that produce 1:1 resonant planet pairs, we see similar characteristics in the semi-major axis evolution. Although it is generally expected that horseshoe orbits are not as stable as tadpole orbits (Dermott and Murray, 1981b), the planet pairs in horseshoe orbits produced in the simulations are found to be stable over the run times we consider. The maximum value of the period ratio among all the co-orbital pairs is ∼1.05:1, in agreement with the simulations by Wu et al. (2018), which show that period ratios in the range 1.05 to 1.1 are unstable (independent of whether or not the system has two planets or a higher multiplicity).

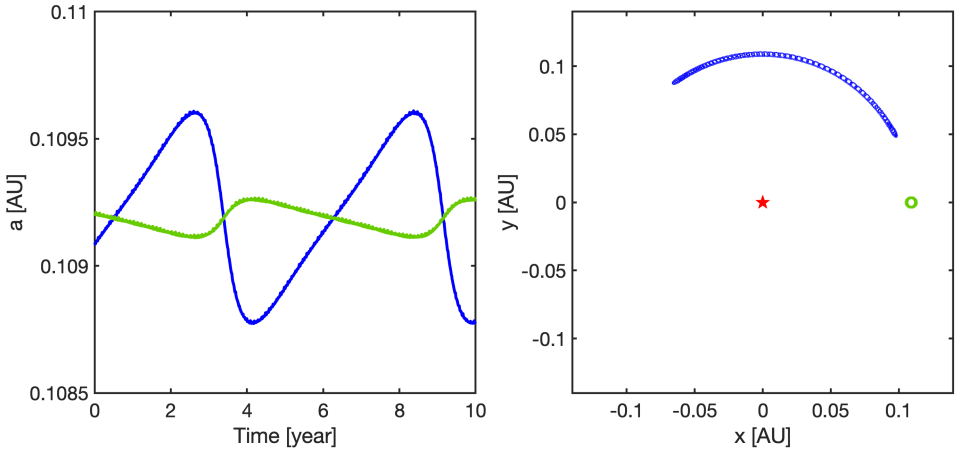

Figure 3.13.: Left panel shows the semi-major axis evolution of the co-orbital planet pair. The right panel shows the orbit trajectory of the same planet pair in a frame that corotates with the planet denoted by the green open circle.

3.4.2. Formation

All the co-orbital planet pairs form fairly early in our simulations (within a few thousand years). Figure 3.14 demonstrates the formation of a co-orbital pair by a 2-body collision event in the perfect collision model. The top panel shows the evolution of the semi-major axis during the first 2,000 yr of the simulation. We can see the collision happened around 1,200 yr (black dashed line). The 3 bodies involved in the encounter are marked by blue, green and purple lines (labelled as planet-b, planet-g and planet-p, respectively, from now on). In this encounter, planet-g and planet-p are the surviving planets and they form a co-orbital tadpole orbit. The second panel of the figure shows the change of mass of the bodies during the evolution. We can see planet-b collided with planet-p during the encounter, and the resulting body ended up with the appropriate energy and angular momentum so that it could settle into a co-orbital configuration with planet-g. The figure also shows the eccentricity and inclination evolution in the bottom two panels. The eccentricity evolution shows the angular momentum and energy exchange (Funk et al., 2011, 2013) between planet-g and planet-p. The fluctuations of these two orbital elements are seen to be reduced after the encounter compared to before it, because of the collision and formation of the co-orbital pair. The two co-orbital planets are mutually inclined by approximately 2°, and hence it cannot be guaranteed that both planets would be detected in a photometric survey searching for transiting planets.

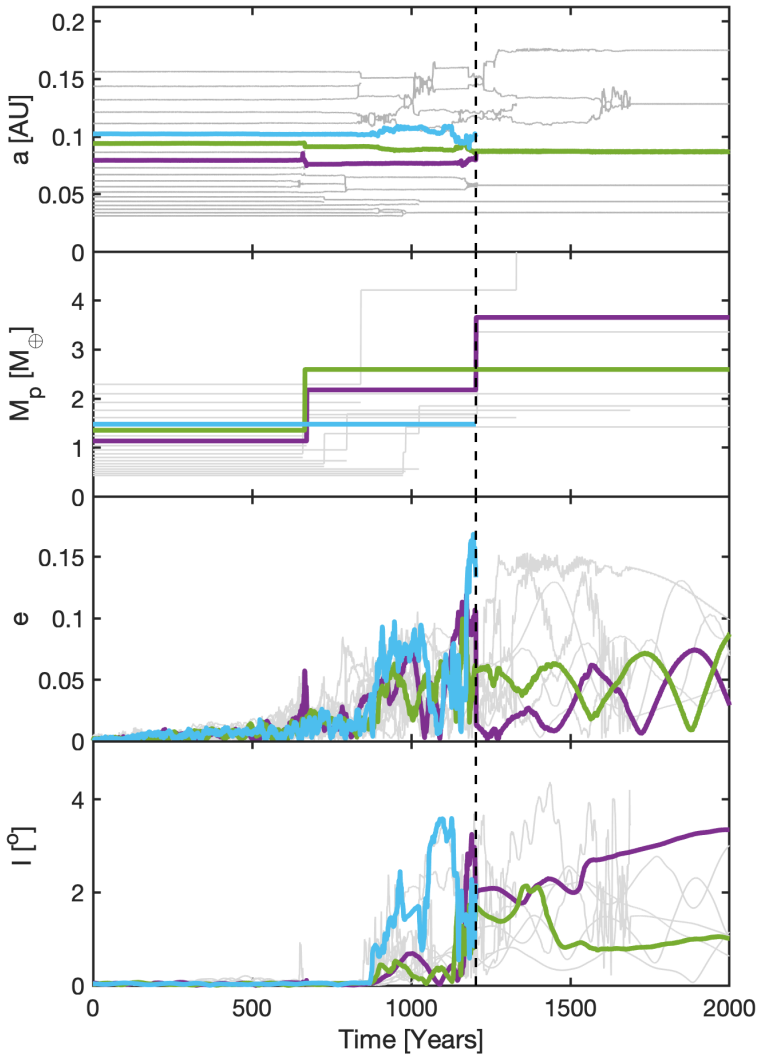

Figure 3.14.: An example of the formation of a co-orbital planet pair by an inelastic collision. Shown are the evolution of the semi-major axes (top panel), planet masses, (second panel) eccentricities (third panel), and inclinations (fourth panel). The time interval shown is the first 2,000 years of the simulation, and the collision occurred after ∼1,200 years (dashed line). The three bodies involved in the encounter are marked in blue, green, and purple, where we name them planet-b, planet-g, and planet-p, respectively.

It is clear that formation of a co-orbital planet pair, involving two planets that were initially well-separated in orbital radius, requires energy and angular momentum loss from one of the planets. Within our simulations, there are three possible ways to achieve this energy loss: 1) an inelastic collision between two bodies resulting in the composite body having the appropriate energy and angular momentum to form a co-orbital pair with a third planet; 2) interaction with collision debris (in the form of multiple planetesimals) formed from an earlier collision, leading to the requisite change in energy and angular momentum by the members of the co-orbital pair; 3) a 3-body encounter in which energy and angular momentum from at least one planet in the co-orbital pair is given to a third body. The case illustrated in figure 3.14 corresponds to the first of these formation scenarios. Neither the interaction with debris nor the 3-body encounter formation mechansims were observed in the simulations, although simulations using the perfect and imperfect accretion routines both resulted in the formation of co-orbital pairs. All systems that formed co-orbital pairs did so early in the simulations, when the space density of planets and the probability of capture due to kinetic energy lost of the colliding planets were at their highest.

3.4.3. Resonance-induced TTV

The co-orbital planet pair (planet-p and planet-g) shown in figure 3.14 survived to the end of the 10 Myr evolution. The final mutual inclination of this planet pair is $\sim 2^o$, which makes it unlikely that both planets would be detected directly during a transit survey. On the other hand, this type of 1:1 resonance pair would induce transit timing variations (TTV) on each other, which might provide a signal indicating the presence of the other non transiting co-orbital planet.

Figure 3.15 demonstrates the TTV signal expected for planet-g during a 10 year period after the end of the simulation. Here, we have calculated the mean orbital period of planet-g over this 10 year period ($P_g \approx 9.9872$ days), and have then computed the Observed–Calculated (O–C) times for the transits of planet-g. The amplitude of the TTV signal reaches ± 0.9639 day, and the maximum difference between adjacent periods is $\sim$ 15 minutes. If it was possible to pick up the transit signal of planet-g and confirm it as a planet within a transit survey, then the TTV signal would provide strong evidence of the presence of the other planet (e.g. planet-p in this case). However, we note that such a strong TTV signal might also provide a barrier to detecting co-orbital planets in the automated pipelines of transit surveys which adopted schemes such as box-least squares with fixed orbital periods, particularly

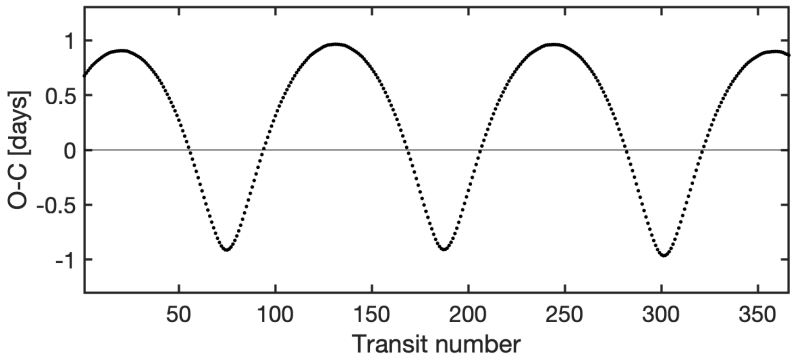

Figure 3.15.: An example synthetic TTV signal for the 1:1 co-orbital planet pair discussed in the text. It shows the difference between the observed and calculated (predicted) times of planet-g's transit events for 10 years after the end of the formation simulation. A total of 366 transit events are observed within this 10 year period.

for systems with low signal to noise.

All 4 of the co-orbital planet pairs formed in our simulations are mutually inclined and lead to a similar situation to that discussed above. So far, no confirmed co-orbital planets have been found. We note that simulations involving dynamical relaxation within a protoplanetary disc also suggest that 1:1 co-orbital planets are a natural outcome (Cresswell and Nelson, 2006, 2008), but in that case the co-orbital pairs are expected to be co-planar and hence would both be detected directly in transit surveys. The strong TTV signal we demonstrated here might provide an explanation of why co-orbital pairs have not been found, and also provide a means of detecting non-coplanar co-orbital systems.

3.5. Collision-induced composition changes

Recent analyses of the distribution of planetary radii for planets discovered by Kepler indicate the presence of a valley in the distribution for radii $1.6 \lesssim R_{\rm p} \lesssim 2$ $R_\oplus$ (Fulton et al., 2017; Fulton and Petigura, 2018; Van Eylen et al., 2018). The location of this valley has been widely interpreted as providing evidence that complete photoevaporation of hydrogen-helium envelopes from core dominated super-Earths has unveiled a population of bodies whose densities are consistent with them having Earth-like compositions (Owen and Wu, 2017; Jin and Mordasini, 2018). The lack of clear evidence for the solid cores having densities consistent with having significant ice fractions suggests that the observed cores did not migrate to their current locations after formation beyond the ice line. An alternative possibility that we

explore here is that high energy collisions during giant impacts may have changed the compositions of previously water-rich cores by stripping off the volatile outer layers.

For head-on collisions, the specific collision energy, $Q_{\rm R}$, can be calculated according to

$$Q_{\rm R} = \frac{1}{2}\frac{\mu V_{\rm imp}^2}{M_{\rm Total}}, \tag{3.20}$$

where the reduced mass $\mu = (M_1 M_2)/(M_1 + M_2)$ (with M_1 and M_2 being the masses of the target and projectile, respectively), $M_{\rm Total}$ is the total mass of the two colliding bodies, and $V_{\rm imp}$ is the relative impact velocity. Stewart and Leinhardt (2009) give a catastrophic disruption threshold energy, $Q^*_{\rm RD}$, which depends on the sizes of the colliding objects and $V_{\rm imp}$. $Q^*_{\rm RD}$ is defined as the energy needed to leave the largest remnant with less than 50% of $M_{\rm Total}$ and is given by the relation

$$Q^*_{RD} = \frac{1}{10^4} R_{C1}^{1.2} V_{\rm imp}^{0.8}, \tag{3.21}$$

where R_{C1} is the radius of a spherical body containing all of the colliding mass with density $\rho_1 = 1$ g cm^{-3}, given by

$$R_{C1} = \sqrt[3]{\frac{3M_{\rm Total}}{4\pi\rho_1}}. \tag{3.22}$$

With the masses of the two colliding bodies known, and the impact velocities measured from the N-body simulations, we can obtain $Q_{\rm R}$ and $Q^*_{\rm RD}$ directly. Marcus et al. (2010), considered collisions involving differentiated bodies, with half of the mass of the colliding planets being water ice and the other half being rock, and found that the mass fraction of the core of the largest remnant, $M_{\rm core}/M_{\rm lr}$, can be fit by the expression

$$\frac{M_{\rm core}}{M_{\rm lr}} = 0.5 + 0.25\left(\frac{Q_{\rm R}}{Q^*_{\rm RD}}\right)^{1.2}. \tag{3.23}$$

This power law of $Q_{\rm R}/Q^*_{\rm RD}$ is the best fit from their smoothed particle hydrodynamics simulations.

The upper left panel of figure 3.16 shows the distribution of the recorded impact angles, θ, in all the giant impact events from our imperfect accretion routine. It is clear that the majority of the collisions are not head-on. For a more realistic analysis for the composition changes, we also consider the effect of off-centre collisions. Leinhardt and Stewart (2012) provide a correction to $Q_{\rm R}$ and $Q^*_{\rm RD}$ that allows them to be applied to off-centre collisions by considering the fraction of the projectile

mass that directly intersects the target during a collision. The mass fraction of M_2 involved in the collision is defined as α (i.e. $\alpha = M_{2,\text{involved}}/M_2$). After defining α, the interacting reduced mass, μ_α, can be calculated by

$$\mu_\alpha = \frac{\alpha M_1 M_2}{\alpha M_2 + M_1}, \tag{3.24}$$

where α can be calculated directly from the information recorded during the simulations by

$$\alpha = \frac{3R_2\left[R_{\text{Total}} - R_{\text{Total}}\sin\theta\right]^2 - \left[R_{\text{Total}} - R_{\text{Total}}\sin\theta\right]^3}{4R_2^3}, \tag{3.25}$$

where $R_{\text{Total}} = R_1 + R_2$, and $\alpha = 1$ when $R_1 - R_2 > R_{\text{Total}}\sin\theta$. The specific impact energy with off-centre collision correction, Q'_{R}, can then be calculated according to

$$Q'_{\text{R}} = \frac{\mu}{\mu_\alpha} Q_{\text{R}}. \tag{3.26}$$

And similarly, the catastrophic disruption threshold energy with off-centre collision correction, Q'^{*}_{RD}, can be calculated by

$$Q'^{*}_{\text{RD}} = \left(\frac{\mu}{\mu_\alpha}\right)^{2-\frac{3\bar{\mu}}{2}} Q^{*}_{\text{RD}}, \tag{3.27}$$

where $\bar{\mu}$ is the velocity exponent in the coupling parameter (Holsapple and Schmidt, 1987; Housen and Holsapple, 1990). Leinhardt and Stewart (2012) and its follow-up study by Stewart and Leinhardt (2012), suggested the range of values of $\bar{\mu}$ is between 0.33 to 0.37. The middle value ($\bar{\mu} = 0.35$) is adopted in equation 3.27 for our calculations. With the new value of Q'_{R} and Q'^{*}_{RD} calculated by equation 3.26 and 3.27, respectively, the mass ratio between the core and the largest remnant from equation 3.23 is modified to

$$\frac{M_{\text{core}}}{M_{\text{lr}}} = 0.5 + 0.25\left(\frac{Q'_{\text{R}}}{Q'^{*}_{\text{RD}}}\right)^{1.2}. \tag{3.28}$$

Our imperfect accretion simulations record all of the data needed to calculate $M_{\text{core}}/M_{\text{lr}}$, and hence determine whether or not the giant impacts occurring in the simulations would have been likely to lead to significant compositional changes if our protoplanets were differentiated bodies consisting of $\sim 50\%$ rock and $\sim 50\%$ water ice, as considered by Marcus et al. (2010). The results of our analysis are shown in

figure 3.16. In the upper left panel, the histogram shows the impact angles, which are peaked at the value around 45°, as expected (Shoemaker, 1962). The lower left panel records the cumulative distribution of $V_{\rm imp}$ in terms of the mutual surface escape velocity, $V_{\rm esc}$, of all collisions experienced across all of the imperfect accretion simulations. The upper right panel shows the cumulative value of $Q'_{\rm R}$, and the lower right panel shows the resulting estimates of $M_{\rm core}/M_{\rm lr}$ arising from each collision, calculated from equations 3.26 and 3.28. The data suggest that only 10% of the giant impacts in our simulations would lead to a greater than 10% mass loss from the protoplanet, where this mass loss would correspond to partial stripping of the putative water-rich mantle. It is very uncommon to have a collision which can cause the protoplanet to have a mass loss of up to 45% (i.e. 90% of the water/ice content), so we conclude that while moderate compositional changes would be likely to have occurred if the Kepler multi-planet systems underwent a final stage of assembly involving giant impacts, the changes would have been insufficient to explain the location of the valley in the distribution of planetary radii discussed above.

3.6. Collision-induced atmospheric loss

Observations and structure models of exoplanets suggest that many of the low and intermediate mass planets observed by Kepler have hydrogen-helium (H/He) envelopes (e.g. Lopez and Fortney, 2014). These H/He envelopes must have been accreted while the planets were embedded in the gaseous protoplanetary disc (e.g. Bodenheimer and Lissauer, 2014; Coleman et al., 2017). If dynamical instabilities and giant impacts after the dispersal of the gas disc have played an important role in the final assembly of the super-Earths and sub-Neptunes observed by Kepler, then the envelopes we observe today must have survived the giant impacts. Previous studies have investigated the conditions under which giant impacts can lead to ejection of an envelope, both by the shock that is driven through the envelope during the impact (Genda and Abe, 2003; Inamdar and Schlichting, 2016; Yalinewich and Schlichting, 2019), and because of the intense heating of the core and envelope that occurs when the impact energy is converted to thermal energy during the collision (Biersteker and Schlichting, 2019).

Following the discussion in Biersteker and Schlichting (2019), we make the simplifying assumption that the kinetic energy associated with an inelastic collision between two bodies is converted efficiently into thermal energy in the planetary core, and good thermal coupling between the core and envelope ensures that the

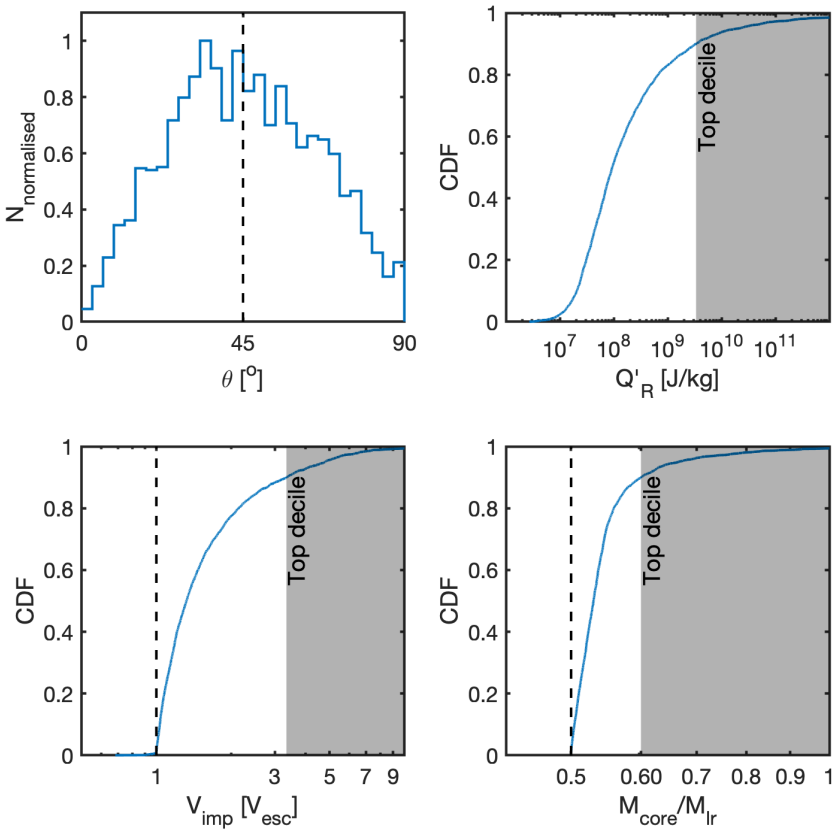

Figure 3.16.: Information concerning collisions occurring across all imperfect accretion simulations. The top left panel shows the distribution of impact angles. The top right panel shows the cumulative distribution of the impact energies, calculated using equation 3.26. The bottom left panel shows the cumulative distribution of the impact velocities in units of the escape velocity, and the bottom right panel shows the cumulative distribution of the ratio of $M_{\rm core}/M_{\rm lr}$, calculated using equation 3.28. The grey area in each plot denotes the region where the cumulative number is within 10% of the maximum.

base of the envelope achieves the same temperature as the core. If this temperature is such that the Bondi radius of the envelope is smaller than the core radius, then we assume that the atmosphere is lost, or is at least severely eroded.

The impact energy, $E_{\rm imp} = 1/2\mu V_{\rm imp}^2$, associated with each collision between two protoplanets is reported by our SYMBA simulations. Hence, in a post-processing step, we can determine the distribution of impact energies from our simulations and determine whether or not these are likely to be sufficient to erode any putative envelopes that the planets might possess. The increase of the temperature, ΔT, due to the impact event can be estimated by equating the impact energy to the change in thermal energy in the core that is present after the impact (which has mass $M_{\rm lr}$):

$$E_{\rm imp} = \eta c_{\rm v} M_{\rm lr} \Delta T, \tag{3.29}$$

giving

$$\Delta T = \eta \frac{1}{2} \frac{\mu}{M_{\rm lr}} \frac{V_{\rm imp}^2}{c_{\rm v}}. \tag{3.30}$$

Here, $c_{\rm v}$ is the specific heat capacity of the core, and η is an energy conversion efficiency factor. The impact should lead to an increase of the final temperature of the core after the impact, $T_{\rm c,final} = T_{\rm c,initial} + \Delta T$, and here we take a conservative approach and assume the initial core temperature is negligible compared to the final value (i.e. $T_{\rm c,initial} \simeq 0$). Assuming the base of the envelope has the same temperature as the core, the associated Bondi radius becomes

$$R_{\rm B} = \frac{2GM_{\rm lr}}{c_s^2} = \frac{2GM_{\rm lr}\mu_m}{\gamma k_{\rm B} T_{\rm c,final}}, \tag{3.31}$$

where c_s is the sound speed, μ_m is the mean molecular weight, γ is the adiabatic index and $k_{\rm B}$ is the Boltzmann constant. We assume the atmosphere is likely lost due to an impact if $R_{\rm B} \leq R_{\rm core}$.

When calculating the value of ΔT and $R_{\rm B}$, we assume the envelopes are a mixture of molecular hydrogen and atomic helium with $\mu_m = 2.3$u and $\gamma = 7/5$. Previous studies have taken values of the specific heat capacity of the cores of super-Earths and sub-Neptunes in the interval $c_{\rm v} = 500 - 1000$ Jkg^{-1}K^{-1} (Alfè et al., 2001; Valencia et al., 2010; Nettelmann et al., 2011; Lopez et al., 2012). In this study, we adopt the middle value within this range $c_{\rm v} = 750$ Jkg^{-1}K^{-1} as in Biersteker and Schlichting (2019).

The left panel in Figure 3.17 shows the cumulative distribution of the post-impact changes in core temperature, ΔT, and the right panel shows the distribution of the

quantity $(R_{\rm B} - R_{\rm core})/R_{\rm core}$, such that a negative value implies substantial erosion of the envelope. From the figure, we can see more than 60% of the collisions in our simulations could lead to envelope loss. These values were obtained by adopting $\eta = 1$ in equation 3.30, corresponding to 100% efficiency in converting impact kinetic energy into heat. Given that not all collisions are head-on, this is clearly an overestimate, as some of the energy can be converted into rotational energy or be taken away by post-impact debris (Agnor and Asphaug, 2004).

Recently, Carter et al. (2020) investigated the Moon-forming impact and showed that around half of the impact kinetic energy is converted to internal energy, such that a more realistic figure would be $0.4 < \eta < 0.6$. It is worth noting that in practice, however, in a five-planet system that was initially composed of 20 protoplanets, each remaining planet after the final assembly stage would have experienced 3 collisions on average, suggesting that significant atmospheric erosion should occur in super-Earth systems whose final assembly involves giant impacts.

Clearly a more sophisticated approach is required to give a better quantitative estimate of the population of planets that are left with significant H/He envelopes after such a period of evolution. Such a calculation would provide one means of determining whether or not the observed population of super-Earths did indeed form via giant impacts. Even if the impacts themselves are unable to completely erode the envelopes, the remnant envelopes will be left in a bloated state and would therefore be more susceptible to photoevaporation by high energy radiation from the central star, as considered in the models of Owen and Wu (2017) and Jin and Mordasini (2018), for example. Hence, in the future it will be important to consider the evolution of envelopes during and after the giant impact phase to determine whether the resulting population of planets agrees with the observations.

3.7. Synthetic observation of the final planetary systems

The ability of any photometric observation of a planetary system to detect transits of all system members depends on the mutual inclinations of the planets. In addition, for any given planet with semi-major axis a, orbiting around a star with radius R_*, the probability of detecting a transit from a random viewing position scales as R_*/a, such that more distant planets around smaller stars are more difficult to detect. Based on these considerations, a meaningful comparison between the outcomes of planetary formation simulations and transit surveys, such as the one carried out by

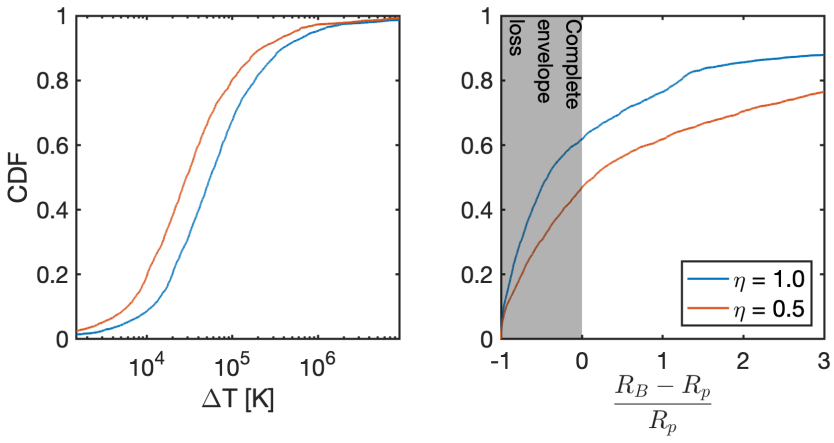

Figure 3.17.: Cumulative distributions of (left panel) ΔT; (right panel) $(R_{\rm B}-R_p)/R_p$ calculated by all the giant impact events in our imperfect collision model simulations. Where $(R_{\rm B} - R_p)/R_p \leq 0$ imply that the Bondi radius is equal to or smaller than the radius of the planet, i.e. complete H/He envelope loss (the grey area). The blue line denotes the value calculated with an energy conversion efficiency of 100% ($\eta = 1$), and red line denotes the value calculated with the energy conversion efficiency of 50% ($\eta = 0.5$).

Kepler, must involve synthetic observation of the simulated planetary systems.

Broadly speaking, the masses, orbital period ratios and planetary separations (as measured by the K-values) resulting from the N-body simulations show reasonable agreement with the inferred properties of the Kepler systems we have used as templates when setting up the initial conditions of the simulations. Here, we are interested in whether or not the distributions of the multiplicities of the simulated planetary systems, and the period ratios between neighbouring planets, when synthetically observed, agree with an appropriate sub-set of the Kepler systems. If such agreement was obtained, then it would support the hypothesis that the observed Kepler systems are all intrinsically high multiplicity systems with mutual inclinations similar to those that arise in the N-body simulations, which in turn would imply that the final assembly of the Kepler systems likely arose from a population of protoplanets that underwent dynamical instabilities and giant impacts, as considered in our N-body simulations. In addition, recent analyses have indicated that planets which are members of multiple systems have a statistically significant different eccentricity distribution compared to planets that are observed to be single (e.g. Xie et al., 2016; Mills et al., 2019). We test whether this difference is matched by our simulations when synthetically observed.

3.7.1. Observed multiplicities

Following the approach of Johansen et al. (2012), we consider the relative numbers of 1-planet, 2-planet, 3-planet, up to 7-planet systems that are detected when the simulation outcomes are synthetically observed from 100,000 randomly chosen viewing locations, isotropically distributed with respect to each host star. Using the observed numbers of 1-planet, 2-planet, 3-planet, etc. systems, we then define a Transit Multiplicity Ratio (abbreviated to TMR hereafter) as follows:

$$\mathrm{TMR}(i,j) = \frac{\text{Number of } i\text{-planet systems}}{\text{Number of } j\text{-planet systems}}, \tag{3.32}$$

where i and j represent the numbers of planets detected during each of the synthetic observations.

For comparison with the TMR values obtained from the N-body simulations, we take a sub-set of the Kepler Planet Candidates with the following cuts applied to the orbital periods, P, and planetary radii, $R_{\rm p}$, so that the Kepler sample roughly matches the simulation outcomes: $3 \leq P \leq 100$ days and $1 \leq R_{\rm p} \leq 4$ $\mathrm{R}_{\oplus}$. In addition, to crudely account for the fact that the detection efficiency of Kepler decreases for small planets with longer orbital periods, we also required the planet to have a radius greater than the value given by $R_{\rm min} = 0.60(P/1\,\mathrm{day})^{0.111}$ $\mathrm{R}_{\oplus}$ when undertaking the synthetic observations. Incorporating this limit on the planet radius excluded around 20% of our final planets, but made essentially no difference to the TMRs obtained from the synthetic transit observations.

The TMR values obtained are shown in figure 3.18, where the coloured histograms show the values obtained from each of the different sets of N-body simulations, the blue horizontal bars show the mean values averaged over the different simulation sets, and the black horizontal bars show the TMRs from the Kepler data. The left panel shows TMRs for 2-planet:1-planet systems, 3-planet:2:-planet systems, 4-planet:3-planet systems, etc., and the right panel shows TMRs for n-planet systems relative to 2-planet systems, where n is an integer running between 3 and 7. The results are very clear: the simulations consistently over produce high multiplicity systems relative to low multiplicity systems by a factor of between 1.5 and 2 compared to the Kepler systems. One reason for this is that the planet-planet scattering, leading to increases in the mutual inclinations of planetary orbits during the N-body simulations, does not increase the mutual inclinations sufficiently for agreement to be reached. The N-body simulation outcomes have mutual inclinations that are too low, with the RMS value obtained from the inclinations plotted in figure 3.8 being $\langle I \rangle = 2.05°$.

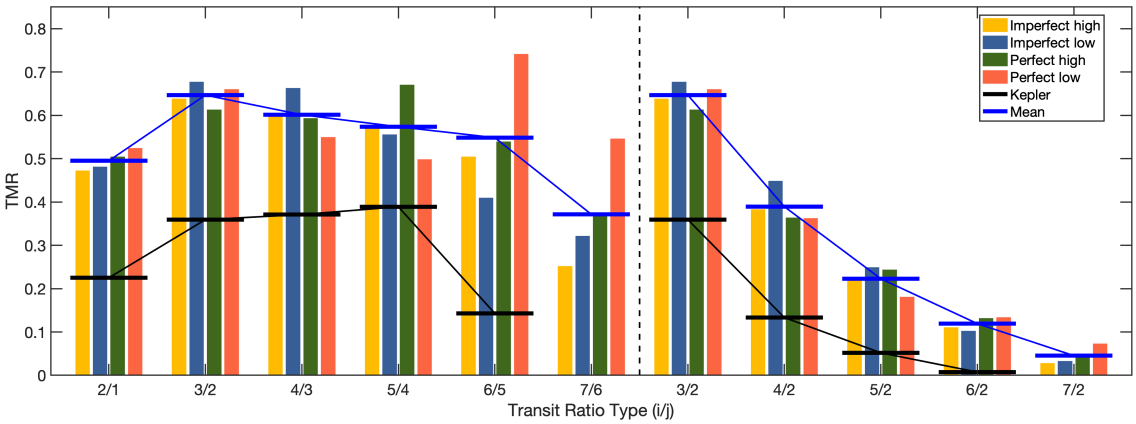

Figure 3.18.: Synthetic transit multiplicity ratio from all four sets of our simulations. Yellow and blue bars denote the high and low initial eccentricity value simulations with imperfect accretion, respectively; green and red bars denote the high and low initial eccentricity simulations with perfect accretion, respectively. Black horizontal lines show the observed Kepler TMRs (as of 17/10/2018) and the blue horizontal lines show the mean values of the simulated TMRs.

In their earlier studies of multiplicity ratios, Johansen et al. (2012) and Tremaine and Dong (2012) suggest that mean mutual inclinations of $\langle I \rangle \simeq 5°$ would be sufficient to provide agreement between their models and the Kepler data when comparing the relative numbers of 3-planet and 2-planet systems. This indicates that a factor of two increase in inclinations in our model systems would likely lead to much better agreement with the Kepler systems, given the factor of ~ 2 discrepancy shown between the TMRs shown in figure 3.18. Even more recently, Izidoro et al. (2019) and Carrera et al. (2019) have presented N-body simulations that provide much better agreement with the Kepler TMRs than our results do. In the case of the Izidoro et al. (2019) study this improved agreement arises in part because they simulate the formation of more massive planetary systems than we do, leading to more effective gravitational scattering, but in addition their simulations result in a number of systems with lower intrinsic multiplicities compared to our simulations.

3.7.2. Period ratios

The CDF of the period ratios between neighbouring planets obtained from the synthetic observation of the simulation outcomes is shown in figure 3.19, along with that for the full set of Kepler planets that make up our comparison sample, and a subset of that sample for which the maximum period ratio is 3:1. This figure should

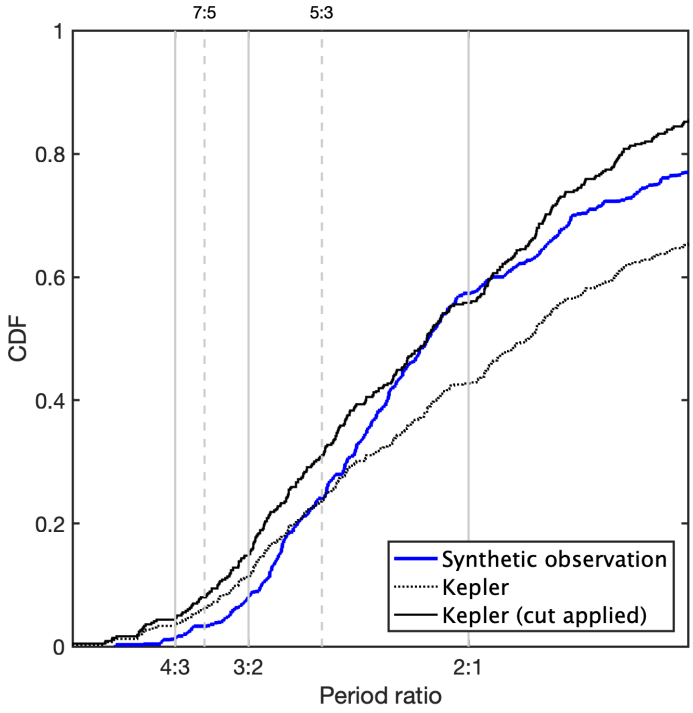

Figure 3.19.: CDFs of the period ratios obtained from the synthetic observations of the simulations, and for comparison the CDFs of period ratios obtained from our comparison sample of the Kepler data.

be compared with figure 3.10, which shows the CDF of the intrinsic period ratios obtained from the simulations before being synthetically observed. This comparison demonstrates the importance of undertaking synthetic observations to mimic transit surveys, as the two distributions of period ratios are quite different from one another. We see from figure 3.19 that we obtain a significant increase in the frequency of period ratios > 2 when undertaking the synthetic observations, as planet pairs on mutually inclined orbits are not observed to simultaneously transit. However, it is also clear that the Kepler data still show a significant excess of large period ratios compared to the simulations, and in general the Kepler systems are more separated than the simulated systems. We also note that we recover the fact that the Kepler systems also show a significant excess of small period ratios compared to the simulations, discussed already in section 3.3.2.

One curious feature of the CDF shown in figure 3.19 is the flattening observed close to the location of the 2:1 resonance, which is reminiscent of the similar feature seen in the Kepler data due to there being a small deficit of planets at the 2:1 resonance location. This is not observed so strongly in the CDF of the intrinsic

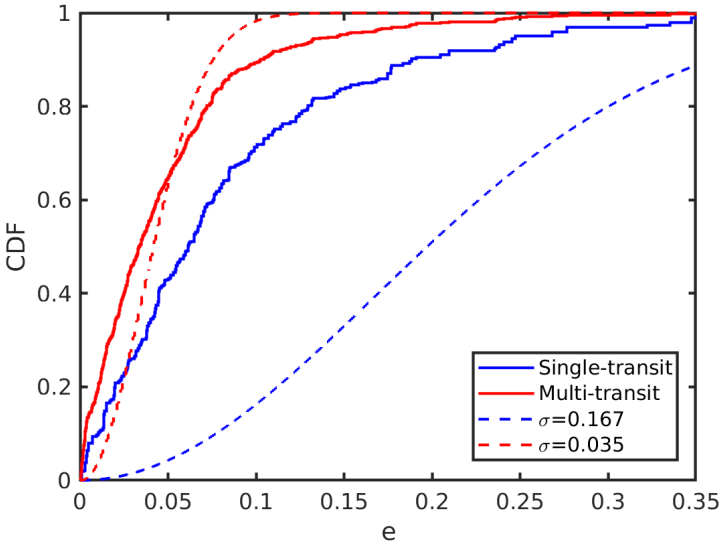

Figure 3.20.: CDFs of the eccentricities obtained from the synthetic observations of the simulations, and for comparison the CDFs of eccentricities drawn from Rayleigh distributions with eccentricity parameters $\sigma_e = 0.035$ and $\sigma_e = 0.167$.

period ratios shown in figure 3.10, so we have examined the possibility that it arises here because mutual inclinations of planet pairs are increased near to this resonance. Plotting mutual inclinations against period ratios, however, showed no significant feature close to the 2:1 resonance, so for now this feature remains unexplained.

3.7.3. Eccentricity distributions

The CDFs of the eccentricities of the synthetically observed planets are shown in figure 3.20, where the solid lines represent either systems observed to be singles or those observed to be multiples. The dashed lines show the CDFs for eccentricities drawn from a Rayleigh distribution with eccentricity parameters $\sigma_e = 0.035$ and $\sigma_e = 0.167$, which are the distributions and values for the observed Kepler multiple and single systems from Mills et al. (2019). While it is clear that the simulations produce single planets with systematically larger eccentricities than the planets in multi-planet systems because the singles are from systems that have undergone stronger scattering than the multiples, it is also clear that the simulations do not provide a good match to the observationally inferred distributions of eccentricities from Mills et al. (2019). In particular, the singles would need to be much more eccentric to match the observationally inferred distribution, and it is not at all clear that N-body simulations of the type presented here could fit the appropriate dis-

tribution of eccentricities, while also adopting planetary masses in line with those thought to make up the Kepler compact systems of super-Earths (such as shown in figure 3.21). On the other hand, although the distribution from multiple systems is not particularly well fitted by the Rayleigh distribution, the range of eccentricities obtained is in much better agreement compared to those obtained for single transiting systems. Using a maximum likelihood estimation, and scanning through different values of σ_e, we find that an *assumed* Rayleigh distribution with parameter $\sigma_e = 0.049$ provides the best fit to the multiple systems arising from the simulations. In future work, we will examine fitting the observed eccentricity distributions with the results of N-body simulations that consider different scenarios to those presented here.

Under conditions of strong scattering, the perturbed radial velocity of a planet relative to a circular Keplerian orbit, v_r, should correspond approximately to the escape velocity from the perturbing body. Assuming typical planet masses and radii $\bar{M}_p$ and $\bar{R}_p$, respectively, we have $v_r \sim \sqrt{2G\bar{M}_p/\bar{R}_p}$. For small eccentricities $e \sim v_r/v_k$, where v_k is the Keplerian velocity. From the CDF of simulated planet masses in figure 3.21, the median planet mass $\bar{M}_p \sim 2.5$ M$_\oplus$ and the corresponding radius $\bar{R}_p \sim 1.56$ R$_\oplus$. For planets orbiting at $a_p \sim 0.1$ au, the eccentricity expected from strong scattering $e \sim 0.15$. The median eccentricity for single planets synthetically observed in the simulations is $e \sim 0.07$, and approximately 20% of planets have eccentricities above the strong scattering value of $e \sim 0.15$. Hence, strong scattering contributes significantly to the eccentricity distribution, but weaker scattering events and collisional damping result in the majority of planets having smaller eccentricities. Finally, it is expected that the mean inclination, $\bar{i} \sim \bar{e}/2$ after dynamical relaxation (Kokubo, 2005). For $\bar{e} \sim 0.07$ the expected mean inclination is $\bar{i} \sim 2°$, very similar to the mean value observed in the simulations, as discussed in section 3.3.2.

Finally, we comment that for strong scattering it is expected that the resulting eccentricities will scale as $e \sim \sqrt{M_p}$. The discrepancy between the median eccentricity of the simulated single planets and the Rayleigh distribution with $\sigma_e = 0.167$, shown in figure 3.20, is about a factor of 3 (~ 0.07 versus ~ 0.2). Hence, to generate this shift would require an increase in the masses of the planets by a factor of ~ 9. The CDFs for the planet masses from the simulations, and those inferred from the Kepler data, are shown in figure 3.21, and there we observe about a factor of 3 discrepancy between the simulated planets and the Kepler planets. Hence, it is not clear at present whether or not strong scattering of planets that appear to be singles in the Kepler data can account for the inferred eccentricity distribution of

these planets as derived by Mills et al. (2019).

3.8. Discussion and conclusions

Recap of simulation set-up

We have presented the results of N-body simulations of *in situ* planetary system formation. These examine whether or not the final assembly of the Kepler compact multi-planet systems, and perhaps the wider population of Kepler planets, can be explained by a scenario in which a large number of orbiting protoplanets experience dynamical instability after the gas disc has dispersed, and accrete through giant impacts, until long term stable systems emerge. Our approach to creating initial conditions was to take eight of the known Kepler 5-planet systems, and to use these as templates for producing systems of 20 protoplanets, whose total mass was the same as the original Kepler systems (under the assumption of a particular mass-radius relation). For each Kepler template, we considered two different distributions of initial eccentricities and inclinations, a 'high set' for which the maximum values $[e_{\rm max}, I_{\rm max}] = [0.02, 0.01]$, and a 'low set' for which $[e_{\rm max}, I_{\rm max}] = [0.002, 0.001]$. For each of these sets we also adopted two different routines for handling collisional growth: a traditional, perfect accretion model that assumes hit-and-stick collisions; an imperfect accretion model which allows for a range of collision outcomes based on the prescriptions of Leinhardt and Stewart (2012). One of our main results is that the simulation outcomes had almost no detectable dependence on the collision model adopted, and this is because the systems we explored did not dynamically excite themselves sufficiently for collisions to be highly disruptive. This is in agreement with the recent study by Mustill et al. (2018). Consideration of more massive planetary systems, however, may lead to outcomes that depend on the collision model as the enhanced gravitational scattering may lead to higher collision velocities.

Recap of main results

All of the N-body simulations resulted in dynamical instability and collisions between protoplanets. The mean time-scale for collisions to occur was approximately 3×10^5 orbital periods, measured at the centre of the annulus containing the protoplanets, indicating that the systems had time for dynamical relaxation to occur during the process of collisional growth. 90% of collisions occurred within 1 Myr, and we ran the simulations for a total of 10 Myr.

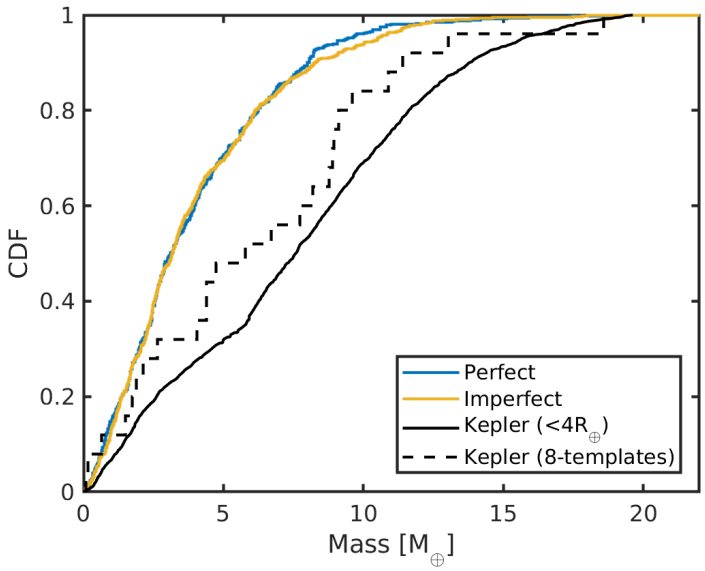

Figure 3.21.: Cumulative distributions of the planet masses obtained from the simulations and inferred from the Kepler data using the mass-radius relation described in the text.

The final outcomes of the simulations generally showed good agreement with the Kepler systems we used as templates, indicating that the procedure adopted for setting up the simulations gave rise to plausible initial conditions. In particular, the final distributions of planet masses, orbital period ratios, separations between neighbouring planets, and intrinsic multiplicities showed good agreement with the templates on average. Notably, our simulations failed to produce any 1-planet or 2-planet systems. This suggests that if single or double planet systems are intrinsically common among the Kepler systems, as has been suggested (Johansen et al., 2012), then the formation scenario presented here cannot explain them. If the single transiting planets are instead members of multi-planet systems whose mutual inclinations prevent all planets being observed, however, then final assembly through planet-planet scattering and giant impacts remains plausible (Carrera et al., 2019; Izidoro et al., 2019). Similarly, widely-spaced pairs of neighbouring planets with large period ratios are very difficult to explain in a model where the initial distribution of protoplanets is smooth and continuous. For these latter systems, it would appear necessary for the initial distributions of planetary building blocks to contain localised concentrations of protoplanets in order to produce the large period ratios seen in the Kepler data. Alternatively, some other process, such as migration, that can cause planets to move relative to one another, would need to be included in the models to explain the well-separated planet pairs that have been observed.

We undertook synthetic transit observations of the final planetary systems formed in the simulations. We counted the relative numbers of 1-planet, 2-planet, 3-planet, ..., 7-planet systems detected by the synthetic observations, and compared this with an appropriate sub-set of the Kepler data that matched the parameters of the model planetary systems. We found that the simulated systems over-produce, by about a factor of 2, the numbers of high multiplicity versus low multiplicity systems compared to the Kepler systems. This arises in part because the excitation of mutual inclinations in our simulations is too small by about a factor of 2, and also because our planetary systems produce too few low-multiplicity systems. We have examined the distributions of the masses of the planets obtained in the simulations, and find that these are smaller than the typical inferred masses of the planets in the Kepler data set (considering planets with radii < 4 $R_\oplus$). Figure 3.21 shows the CDFs of the planet masses from the simulations and the Kepler planets, and it shows: (i) the Kepler systems we chose as templates have moderately smaller masses than the Kepler data set as a whole we are comparing against; (ii) the simulations produce too many low mass planets compared to both the templates and the Kepler systems as a whole, because systems with higher multiplicity than 5 are formed. It is likely that by choosing more massive discs of protoplanets as initial conditions, the resulting planet masses, and the enhanced scattering they would experience, would lead to final systems more in agreement with the Kepler data, both in terms of inferred planet masses and in terms of the distribution of multiplicities, because of the larger mutual inclinations that would have been excited. This may also result in eccentricity distributions that are in better agreement with the observations than we obtained in the simulations presented here. Alternatively, if the final period of dynamical instability in multiplanet systems is initiated when the planets are essentially fully formed, rather than when the system consists of numerous low mass protoplanets, then it may be possible to achieve higher eccentricities and mutual inclinations because the scattering may be stronger and collisions may occur less frequently.

Recent relevant planet formation studies

In their recent study, Izidoro et al. (2019) were able to construct a population of planetary systems, that when synthetically observed, provided good agreement with the Kepler multiplicity distribution. This was achieved by combining simulations that resulted in resonant chains of planets that became dynamically unstable with simulations in which the resonant chains remained stable. It is noteworthy that

only 5% of the included planetary systems were intrinsically 1-planet systems, with most of the rest being multiple systems in which the mutual inclinations typically exceeded 4°. In an earlier study, Moriarty and Ballard (2016) undertook a study of the multiplicity distributions arising from N-body simulations of planet formation, adopting a range of surface density profiles and masses in their initial discs of protoplanets and planetesimals. As with our simulations, theirs also formed planetary systems with intrinsic multiplicities $N_{\rm p} \geq 3$, and by suitably combining their different simulation results they were able to produce a population of planets that agreed with the Kepler distribution of multiplicities when their simulated systems were synthetically observed. Hence, it appears that combining a range of initial conditions for planet formation simulations, that ultimately result in dynamical instabilities and giant impacts, can lead to systems that collectively provide mutual inclination and intrinsic multiplicities that agree with observations when their transits are simulated.

Intrinsic multiplicities from RV studies

The agreement between these simulations and the observations raises an important question: are essentially all planetary systems intrinsically multiple systems, even when observed to be singles by transit surveys? And if so, what is the underlying multiplicity distribution? Transit surveys cannot directly answer this. The detection of transit timing variations in apparently single planet systems discovered by Kepler, however, shows that a number of these planets have neighbours close to mean motion resonances (Kane et al., 2019). In addition, the fact that the eccentricities of single planets appears to be systematically higher than in multiple planet systems (Mills et al., 2019), indicates that a number of apparently single planets have likely been subject to gravitational scattering, and hence are members of multi-planet systems. Radial velocity surveys can in principle detect nearby companions in compact multi-planet systems, assuming modest mutual inclinations, although they are constrained by limits imposed by spectral resolution, instrument stability and stellar variability, and numerous compact multi-planet systems have been discovered by this method (e.g. Mayor et al., 2011). Nonetheless, there are hints in the data that super-Earths do not always come as members of compact multiple systems. For example, the recently discovered super-Earth orbiting with a period of 233 days around Barnard's star indicates a lack of close orbiting planets with similar masses in that system (Ribas et al., 2018), and certainly none that became anchored at the inner edge of the protoplanetary disc during their formation, as often occurs in N-

body simulations of planet formation that involve pebble drift or planet migration. Similarly, Proxima b, orbiting with a period of 11 days does not appear to have closely neighbouring planets of similar masses (Anglada-Escudé et al., 2016). While these are only individual examples, they indicate that not all planetary systems are compact multiples. Future high precision RV surveys targeted at characterising the multiplicities of short period super-Earth systems will have the power to determine the whether or not final assembly of planetary systems via dynamical instability is the dominant mode of planet formation, or if instead a substantial population of relatively isolated planets exists that cannot be explained by the giant impact formation route.

Compact non-resonant systems

Furthermore, compact systems of super-Earths, such as Kepler-11 (Mahajan and Wu, 2014), that contain planets on low eccentricity orbits, and which are apparently close to instability, are also difficult to assemble via dynamical instabilities and giant impacts. Kepler-11 appears to have been assembled in a highly dissipative environment, presumably in a gaseous protoplanetary disc – which is supported the low densities of some of the planets (Lissauer et al., 2013), but none of the planet pairs are in mean motion resonance. Hence, while dynamically quiet formation in a disc seems necessary, the lack of resonances indicates that disc-driven migration may not have played an important role in this system. We note, however, that recent analyses of single and multiple planet migration in inviscid protoplanetary discs by McNally et al. (2019b) and McNally et al. (2019a), lead to more complex migration behaviour of planets than has been found to traditionally occur in viscous discs. Hence, the formation of resonant chains is not a foregone conclusion in inviscid discs, and such an environment may provide a way of forming systems such as Kepler-11. A reasonable conclusion is that the observational evidence appears to indicate that a number of different pathways are required for the final assembly of planetary systems.

Co-orbital systems

Approximately 1% of our simulations gave rise to pairs of planets in apparently long-term stable 1:1 co-orbital resonances, occupying both tadpole and horseshoe orbits. These normally form early in the simulations, when the numbers of protoplanets are high and the planetary systems are undergoing strong planet-planet interactions.

The co-orbital pairs arise as a result of three body encounters removing the requisite energy and angular momentum from a pair of planets such that the co-orbital configuration can form. These co-orbital planet pairs occur with equal frequency in the perfect and imperfect collision model simulations, indicating that the treatment of collisions has no effect on their formation. In spite of intensive searches through the Kepler data, no co-orbital planet systems have been found. We note, however, that there are just over one hundred Kepler systems with known multiplicity ≥ 3 that fall within the parameter ranges covered by our simulations, so a 1% occurrence rate, that would agree with the simulations results, leads to an expectation that just one co-orbital system would have been found. Hence, the current data set is too small to determine if the observed planets indeed formed from a large number of protoplanets undergoing dynamical relaxation and collisions, leading to co-orbital pairs forming with an efficiency of $\sim 1\%$ per system. Future missions, such as PLATO, will monitor many more stars than Kepler (Rauer et al., 2014), and hence will place more stringent constraints on the formation histories of compact multiplanet systems.

Composition changes and envelope loss

Recent observations have indicated the presence of a valley in the distribution of planet radii for short period Kepler planets (Fulton et al., 2017; Fulton and Petigura, 2018; Van Eylen et al., 2018), and models of envelope photoevaporation suggest that the position of the valley is most easily explained if the cores of super-Earths that are subject to photoevaporation are rocky rather than volatile rich (Owen and Wu, 2017; Jin and Mordasini, 2018). In this context, we examined whether or not high energy collisions during our simulations could significantly modify the compositions of the final planets relative to the initial protoplanets. In post-processing, we used the scaling relations between collision energies and compositional changes presented by Marcus et al. (2010), and examined whether or not high energy collisions occurred frequently enough that they could remove a large fraction of water-rich mantles of colliding, differentiated protoplanets whose initial compositions were 50% rock and 50% water-ice. The results of this analysis suggest that collisional stripping of water-rich mantles cannot explain the fact that the apparently naked cores observed by Kepler have an Earth-like composition, instead of a mixture of rock and water-ice. It seems unlikely, therefore, that these now naked cores formed exterior to the ice line and migrated to their current locations. Using a similar analysis, we also examined whether or not the impact energies of collisions between protoplanets could be

sufficient to remove any H/He envelopes they might possess. Our simple analysis suggests that $\sim 30\%$ of collisions occurring in the simulations could remove gaseous envelopes, such that planets experiencing multiple collisions would have a high probability of losing their envelopes completely. Further modelling of this process could place significant constraints on the collisional history of the observed population of super-Earths and sub-Neptunes. Compared to the scenario examined here in which a large number of protoplanets undergo collisional accretion during the final assembly of exoplanet systems, the exoplanet data likely point to an origin in which systems of fewer, essentially fully formed planets undergo dynamical relaxation after dispersal of the gas disc in order to produce the observed orbital architectures, while undergoing fewer collisions in order to maintain the gaseous envelopes possessed by a large number of the observed exoplanets.

In future work we will present simulations with a significantly broader range of initial conditions, to assess the conditions under which dynamical instabilities in multi-planet systems may have contributed to the final stage assembly of the observed short-period super-Earths and sub-Neptunes.

4. Dynamical heating by cold giants

從來分開兩面高梯
隔著妳我不只二十米
熟練的開始手勢
我盪向妳接觸範圍

— 李克勤 ⟨空中飛人⟩

This chapter mainly follows the material that is contained in Poon and Nelson (2020). The title of this publication is '*On the origin of the eccentricity dichotomy displayed by compact super-Earths: dynamical heating by cold giants*', and it was published in volume 498 of the Monthly Notices of the Royal Astronomical Society.

Approximately half of the planets discovered by NASA's Kepler mission are in systems where just a single planet transits its host star, and the remaining planets are observed to be in multiplanet systems. Recent analyses have reported a dichotomy in the eccentricity distribution displayed by systems where a single planet transits compared with that displayed by the multiplanet systems. Using N-body simulations, we examine the hypothesis that this dichotomy has arisen because inner systems of super-Earths are frequently accompanied by outer systems of giant planets that can become dynamically unstable and perturb the inner systems. Our initial conditions are constructed using a subset of the known Kepler five-planet systems as templates for the inner systems, and systems of outer giant planets with masses between those of Neptune and Saturn that are centred on orbital radii $2 \leq a_{\rm p} \leq 10$ au. The parameters of the outer systems are chosen so that they are always below an assumed radial velocity detection threshold of 3 ms^{-1}.

The results show an inverse relation between the mean eccentricities and the multiplicities of the systems. Performing synthetic transit observation of the final systems reveals dichotomies in both the eccentricity and multiplicity distributions that are close to being in agreement with the Kepler data. Hence, understanding the observed orbital and physical properties of the compact systems of super-Earths discovered by Kepler may require holistic modelling that couples the dynamics of

both inner and outer systems of planets during and after the epoch of formation.

4.1. Introduction

At the time of writing more than 4000 exoplanets have been discovered and confirmed using various detection methods. More than half of them were discovered by the Kepler transit survey (Borucki et al., 2010, 2011; Batalha et al., 2013; Burke et al., 2014; Rowe et al., 2015; Mullally et al., 2015; Coughlin et al., 2016; Thompson et al., 2018). Due to detection biases, most of the Kepler planets have short periods < 100 days, and less than one percent have orbital periods longer than one year[1]. The majority of the Kepler planets have radii between 1 and 4 $R_\oplus$, and these super-Earths and sub-Neptunes are often found in compact multi-planet systems.

Various analyses of the Kepler data have been undertaken to obtain insight into the formation and evolution of planetary systems. For example, the distribution of the observed planetary system multiplicities shows a sharp increase for single planet systems compared to two planet systems (Lissauer et al., 2011b), and this apparent *Kepler dichotomy* has been interpreted as either arising because of an intrinsic excess of single planet systems (Johansen et al., 2012), or alternatively because of the distribution of mutual inclinations within multi-planet systems (Zhu et al., 2018). Understanding the origin and nature of this dichotomy would clearly shed light on the history of formation and dynamical evolution experienced by the compact Kepler systems. An apparent dichotomy has also been detected in the distributions of the orbital eccentricities associated with either single or multiple planet systems (Xie et al., 2016; Van Eylen et al., 2019; Mills et al., 2019). The Kepler single-planet systems have a mean eccentricity $\langle e_1 \rangle \approx 0.25$–$0.3$, whereas the multi-planet systems have $\langle e_{\geq 2} \rangle \approx 0.05$, indicating the two populations have experienced different dynamical histories.

Different formation and evolution scenarios have been proposed to explain the Kepler compact multi-planet systems. A study of the architectures of multi-planet systems by Pu and Wu (2015) showed that the high multiplicity systems are close to being dynamically unstable, and these authors suggested that the low multiplicity systems may have experienced dynamical instabilities and planet-planet collisions, leading to the low numbers of planets now observed in these systems. Dynamical instabilities during the late stages of formation can lead to the self-excitation of eccentricities and mutual inclinations within compact planetary systems, and

[1] All exoplanetary data used in this chapter are from NASA Exoplanet Archive unless stated otherwise.

numerous studies of this process have been undertaken (e.g. Hansen and Murray, 2012; Moriarty and Ballard, 2016; Matsumoto and Kokubo, 2017; Izidoro et al., 2017; Poon et al., 2020, and chapter 3). The results of these studies are generally in agreement with the data on the mutual separations between the planets, but when planetary system masses are adopted that are characteristic of those inferred for the Kepler multi-systems, then the degree of gravitational scattering is insufficient to provide large enough mutual inclinations or eccentricities to explain the multiplicity and eccentricity dichotomies described above (e.g. Poon et al., 2020, and chapter 3).

This has led to an alternative hypothesis for explaining the eccentricity dichotomy, namely that giant planets in the outer regions of planetary systems perturb the inner systems. Observations show that compact systems of super-Earths can have outer giant companions. One example is the Kepler-68 system, which contains two transiting super-Earths/sub-Neptunes (Kepler-68b and c) with orbital periods less than 10 days, and a third planet (Kepler-68d) that was discovered using the radial velocity (RV) technique orbiting beyond 1 au (Gilliland et al., 2013; Marcy et al., 2014). In general, cold gas giant planets are on eccentric orbits, and it is well known that planet-planet scattering within a system of gas giants can excite eccentricities to high values ($e > 0.3$) (Chatterjee et al., 2008; Jurić and Tremaine, 2008; Petrovich et al., 2014). If a chain of excited outer giants coexists with an inner compact system, then perturbations can excite the eccentricities and mutual inclinations of the inner planets (Matsumura et al., 2013; Becker and Adams, 2017; Hansen, 2017; Mustill et al., 2017; Lai and Pu, 2017; Huang et al., 2017; Pu and Lai, 2018; Zhu and Wu, 2018; Masuda et al., 2020).

In this chapter we explore this idea using N-body simulations, similar to previous work (e.g. Mustill et al., 2017; Huang et al., 2017), except we explore the effects of adopting a range of multiplicities for the systems of cold giant planets, and we choose parameters for the cold giants such that they would be undetectable in RV surveys that have a detection limit of $v_{\rm RV} = 3$ ms^{-1}. We construct initial conditions that consist of inner systems of super-Earths, based on known 5-planet systems observed by Kepler, and outer systems of giant planets for which the masses are in the range $15 \leq m_{\rm p} \leq 100$ M$_\oplus$ and the semi-major axes are centred between $2 \leq a_{\rm p} \leq 10$ au. The results of the simulations are processed through a pipeline that synthetically observes the systems using the transit method and produces statistics on the multiplicity and eccentricity distributions for comparison with the Kepler data. Using these tools, we identify regions of parameter space that produce results that are close to being in agreement with the data.

This chapter is structured as follows. In Section 4.2 we discuss the selection of

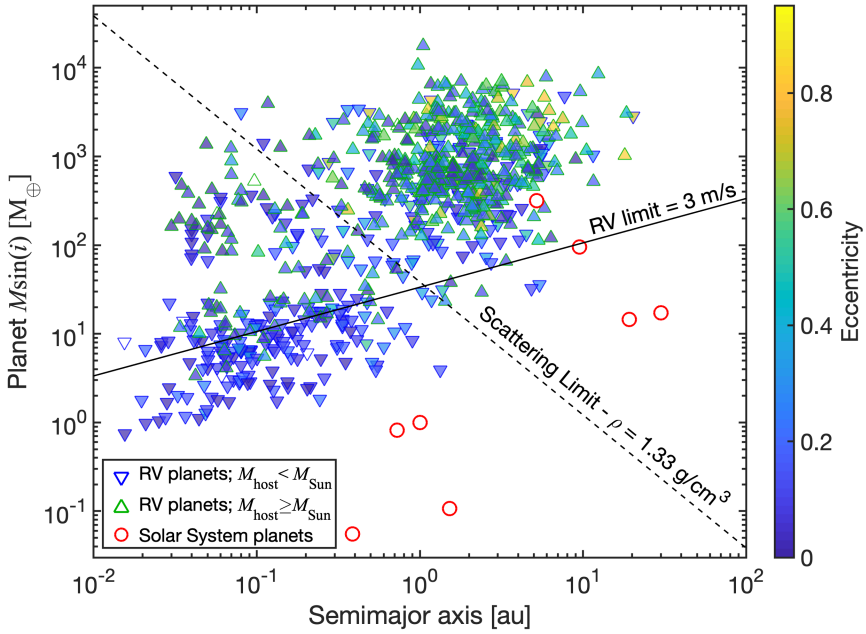

Figure 4.1.: Masses ($M \sin i$) and semi-major axes of confirmed RV planets (triangles), where planets with host star masses $M_{\rm host} < 1$ $M_\odot$ are marked with downward pointing triangles, and planets with $M_{\rm host} \geq 1$ $M_\oplus$ are marked with upward pointing triangles. The colours of the triangles indicate the eccentricities of the planets (they are left unfilled if the eccentricity is not known). The solid line shows the RV detection limit (equation 4.1) of $v_{\rm RV} = 3$ ms^{-1} induced by a planet on an edge-on circular orbit around a Solar type star. The majority of RV planets below the solid line orbit a host star with $M_{\rm host} < 1$ $M_\odot$, making detection possible. The dashed line is the scattering limit for a planet given by equation 4.2, assuming a Jupiter-like mean density and Solar type host star. Solar System planets are marked using red circles for reference.

outer planetary systems for this study, and in Section 4.3 we describe the set-up of the initial conditions and the selection of the inner planetary system templates. In Section 4.4 we present the results of our simulations, and in Section 4.5 we present the synthetic observations of the simulated systems and examine the distributions of multiplicities and eccentricities that arise. In Section 4.6 we discuss the influence of physical effects such as a realistic collision model and tidal interactions that were omitted from our primary suite of simulations, and which could potentially affect the final results when comparing to the observations. Finally, we discuss our results and draw conclusions in Section 4.7.

4.2. Consideration of system selection

Kepler was inefficient at detecting planets with periods greater than 1 year ($a_{\rm p} \gtrsim$ 1 au) because of detection biases and the mission lifetime. Most of what we know about longer period planets comes from RV surveys. Figure 4.1 shows the masses versus semi-major axes for planets discovered by the RV method, demonstrating the existence of a population of Jovian planets with semi-major axes greater than 1 au. The RV survey presented by Mayor et al. (2011) indicates that approximately 14% of Solar type stars host a giant planet with orbital period $\lesssim$ 4000 days, whereas it has been estimated that up to 50% of stars host a super-Earth or sub-Neptune with orbital period $\leq$ 100 days (Fressin et al., 2013). Hence it appears that detectable cold giant planets occur less frequently than the warm super-Earths. In this work we are working with the implicit hypothesis that each system of inner planets may be accompanied by a system of outer planets, and hence we need to consider what that unseen population might look like.

Previous studies of inner systems being perturbed by outer planets focused on the RV-like planets, for which cold Jupiters are massive enough to generate a detectable RV signal, $v_{\rm RV}$. The value of $v_{\rm RV}$ is given by

$$v_{\rm RV} = \sqrt{\frac{G}{1-e^2}}\,(m_{\rm p}\sin i_{\rm p})\,\frac{1}{\sqrt{a_{\rm p}(m_{\rm p}+M_\star)}}, \tag{4.1}$$

where G is the gravitational constant, $m_{\rm p}$ is the mass of the RV planet, $i_{\rm p}$ is the inclination of the planet's orbital plane, $a_{\rm p}$ is the semi-major axis, and $M_\star$ is the mass of the host star. We adopt an RV detection limit of $v_{\rm RV} = 3$ ms^{-1}, and this limit is indicated in figure 4.1 for an edge-on circular orbit ($i_{\rm p} = 90°$; $e = 0$) around a Solar type star ($M_\star = 1$ M$_\odot$). We see that only Jupiter in the solar system lies above the detection limit while Saturn is just below the limit.

For outer system planets to be able to induce perturbations on the orbits of the inner system planets, we need to be in the regime where strong planet-planet scattering is favoured over planet-planet collisions. Under the conditions of strong scattering, the velocity kick experienced by a planet relative to a circular Keplerian orbit should correspond approximately to the escape velocity, $v_{\rm e}$, from the perturbing body (e.g., Ford and Rasio, 2008). Strong scattering can arise when $v_{\rm e}$ is larger than the Keplerian velocity, $v_{\rm K}$ (i.e. $v_{\rm e}/v_{\rm K} > 1$). The mass of the planet corresponding to the ratio $v_{\rm e}/v_{\rm K}$ is

$$m_{\rm p} = \left(\frac{3}{32\pi}\frac{M_\star^3}{a_{\rm p}^3\rho_{\rm p}}\right)^{\frac{1}{2}}\left(\frac{v_{\rm e}}{v_{\rm K}}\right)^3, \tag{4.2}$$

where $\rho_{\rm p}$ is the internal mean density of the planet. The dashed line plotted in figure 4.1 shows the scattering limit for planets orbiting a Solar mass star, assuming $\rho_{\rm p} = \rho_{\rm Jupiter} \approx 1.33$ g cm^{-3}, where planets above the line are more likely to induce scattering while the planets below the line are more likely to collide during orbit crossing.

Combining our requirement that the perturbing planets would be undetectable with long term RV surveys, with the need for them to be in the scattering rather than collision regime, means that we need to select bodies that lie in the region bounded by the solid and dashed lines exterior to 1 au in figure 4.1.

4.3. Simulation set-up

We use the N-body code MERCURY (Chambers, 1999) to undertake the simulations presented in this chapter. Collisions during the simulations are treated using a simple hit-and-stick approach that conserves the total mass and linear momentum when two bodies collide and accrete to form a single object. We have, however, re-run a small subset of simulations to investigate the effects of adopting a more realistic collision model, and tidal interactions with the central star, as these effects were not included in our main suite of simulations. For the initial conditions, we generate templates for the inner and outer planetary systems, and combine them to generate each model, as described below.

4.3.1. Inner planetary system templates

We have selected a number of Kepler five-planet systems as templates for the inner systems using the following criteria. Similar to Poon et al. (2020) and chapter 3, we have selected those Kepler systems where all five of the known planets are transiting. The Kepler-82 and Kepler-122 systems are excluded because Kepler-82f (Freudenthal et al., 2019) and Kepler-122f (Hadden and Lithwick, 2014) were discovered by transit timing variations (TTVs).

We are interested in compact systems of super-Earths, so we have chosen systems in which the known outermost planet has semi-major axis ≤ 1.0 au or orbital period ≤ 1 yr. The Kepler-150 system is excluded as the outermost planet, Kepler-150f, has an orbital period of ~ 1.74 yr (Schmitt et al., 2017). Kepler-444 system is also excluded because the planets are too small to be considered as super-Earths (Campante et al., 2015).

We adopt the mass-radius ($M_{\rm p} - R_{\rm p}$) relation $M_{\rm p} = R_{\rm p}^{2.06}$ suggested by Lissauer

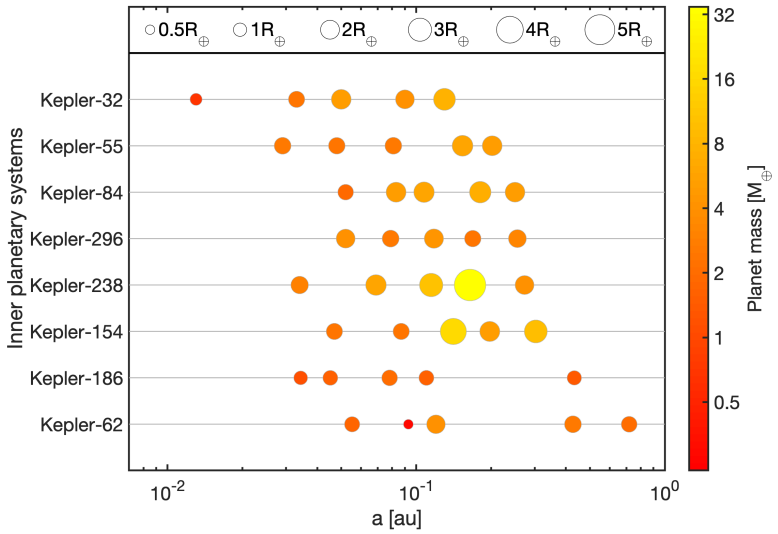

Figure 4.2.: The eight selected inner planetary system templates. The symbol size represents the radius of each planet and the colour represents the planet mass calculated using the mass-radius relation $M_{\rm p} = R_{\rm p}^{2.06}$ adopted from Lissauer et al. (2011b).

et al. (2011b) for the inner planets, where $M_{\rm p}$ and $R_{\rm p}$ are in Earth units. Combining this mass-radius relation with the initial values of the eccentricities and inclinations we adopt (see section 4.3.3), leads to some of the Kepler systems becoming unstable, with at least one planet experiencing a collision during the simulation. Section 4.4.1 describes the results of the stability tests. After excluding the unstable Kepler systems, eight were selected as the inner planetary system templates. They are Kepler-32, -55, -62, -84, -154, -186, -238, and -296 (see figure 4.2).

4.3.2. Outer planetary system templates

The masses of the outer planets we adopted lie between 15 $M_\oplus$ and 100 $M_\oplus$. In contrast to previous studies of the effects of outer giant planets on inner systems (e.g. Mustill et al., 2015; Huang et al., 2017; Mustill et al., 2017), our outer giant planets are below the RV detection threshold discussed above.

The number of outer planets, $N_{\rm out}$, in each simulation is one of 3, 6, or 12. Four different sets of mass are adopted, where $M_{\rm p,out} = 15$, 30, 60 or 100 $M_\oplus$. In a given simulation, all the outer planets start with the same mass. The mutual separation between each giant planet pair, measured in units of the mutual Hill radius, is $K = 4$.

The mutual Hill radius for a pair of adjacent planets is defined by

$$R_{\mathrm{H},i} = \frac{a_i + a_{i+1}}{2} \left(\frac{M_{\mathrm{p},i} + M_{\mathrm{p},i+1}}{3M_\star} \right)^{\frac{1}{3}}, \tag{4.3}$$

where a_i is the semi-major axis of the i^{th} planet in the system. The dimensionless number K is

$$K = \frac{a_{i+1} - a_i}{R_{\mathrm{H},i}}. \tag{4.4}$$

The value of $K = 4$ was chosen as a compromise between ensuring an instability actually occurs during the simulations, having a relatively short instability timescale to make the simulations tractable, and also wanting the instability timescale to not be so short that the outer planetary systems disintegrate within the first few orbits of the simulations. We note a similar value has been used in previous studies (Chatterjee et al., 2008; Petrovich et al., 2014; Mustill et al., 2017).

Three different sets of locations for the outer planets are considered, with the median semi-major axis for the outer planets, $\tilde{a}_{\mathrm{out}}$, being 2, 5, or 10 au. The physical radii of the outer planets, important for determining when collisions occur, are calculated by defining their mean internal densities to be equal to that of Jupiter (1.33 gcm^{-3}).

In summary, the parameters for the outer planetary system templates are comprised of three different multiplicities, four different masses and three different values of the median semi-major axis, giving a total of 36 templates that could be generated for each inner planetary system template. We use a labelling convention based on the parameters of the outer planet system when describing the runs as follows: '(N_{out})g.$(M_{\mathrm{p,out}})$M.$(\tilde{a}_{\mathrm{out}})$AU'. For example, `6g.30M.10AU` refers to the outer planet template with six 30 $\mathrm{M}_\oplus$ planets and a median semi-major axis of 10 au. Table 4.1 lists all the outer planetary templates we have considered in this study. Figure 4.3 displays the semi-major axis distribution of the outer planetary templates.

4.3.3. Constructing initial conditions for the simulations

To investigate the effects of the outer ice/gas giant planets on the inner super-Earths, we generate different initial conditions by combining one inner planetary system template and one outer planetary system template. We label this system as 'inner-planetary-template.outer-planetary-template'. For example, the system template of `Kepler55.6g.30M.10AU` refers to a run which uses Kepler-55 as the template for the inner system and `6g.30M.10AU` for the outer system. To provide a

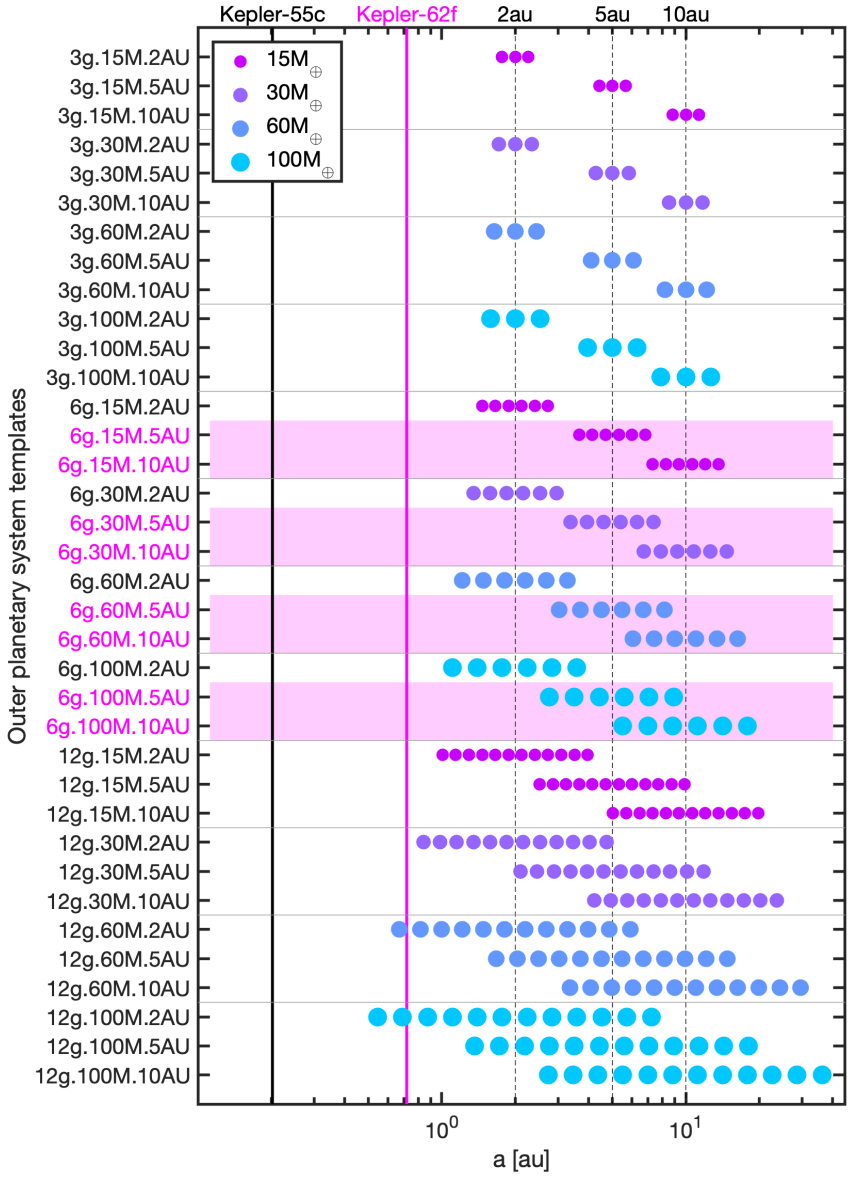

Figure 4.3.: Semi-major axes of the outer planets within different templates (as listed in table 4.1). Different planet masses are marked in different sizes and colours. The eight outer templates that are the focus of most of our runs (see section 4.4.3) are highlighted in pink. The black solid vertical line marks the semi-major axis of Kepler-55c, which is the outermost planet of the Kepler-55 system. The pink solid vertical line marks the semi-major axis of Kepler-62f, which is the outermost planet within our eight selected inner system templates. Kepler-55 is the system we selected to study with all outer planetary templates, while Kepler-62 (and other selected inner planetary templates) is only combined with the eight outer planetary templates described in the text. There is no immediate dynamical interaction between our inner and outer planetary systems.

Table 4.1.: The 36 different outer planet templates considered in the study. The name of each template follows the convention '$N_{\rm out}$g.$M_{\rm p,out}$M.$\tilde{a}_{\rm out}$AU'. The eight templates that are the focus of most of our runs (see section 4.4.3) are indicated by the star superscript.

3 giants	6 giants	12 giants
3g.15M.2AU	6g.15M.2AU	12g.15M.2AU
3g.15M.5AU	6g.15M.5AU*	12g.15M.5AU
3g.15M.10AU	6g.15M.10AU*	12g.15M.10AU
3g.30M.2AU	6g.30M.2AU	12g.30M.2AU
3g.30M.5AU	6g.30M.5AU*	12g.30M.5AU
3g.30M.10AU	6g.30M.10AU*	12g.30M.10AU
3g.60M.2AU	6g.60M.2AU	12g.60M.2AU
3g.60M.5AU	6g.60M.5AU*	12g.60M.5AU
3g.60M.10AU	6g.60M.10AU*	12g.60M.10AU
3g.100M.2AU	6g.100M.2AU	12g.100M.2AU
3g.100M.5AU	6g.100M.5AU*	12g.100M.5AU
3g.100M.10AU	6g.100M.10AU*	12g.100M.10AU

statistically meaningful sample, our aim is to run each of these systems 100 times with different random initialisations of the orbital elements. This would involve running 100 simulation for all the combined templates (8 Kepler 5-planet systems $\times$ 36 outer ice/gas giant templates $\times$ 100 runs = 28,800 N-body simulations in total), which is not possible given our available computational resources. To make the problem tractable, after undertaking a survey of how all the outer system templates affect one of the inner system templates (Kepler-55), we have selected eight outer planetary templates to focus on in detail as described in section 4.4.2.

Each template is run with 100 different instances of the initial conditions. The initial eccentricities, e, of the inner super-Earths are randomly drawn from a Rayleigh distribution with eccentricity parameter, $\sigma_{\rm e} = 0.035$. The value of 0.035 is taken from the analysis of the Kepler systems presented by Mills et al. (2019). The initial inclinations, I, for each run are randomly drawn from a Rayleigh distribution with inclination parameter, $\sigma_{\rm I} = 0.5\sigma_{\rm e} = 0.0175$ radians. For the outer planets, the values of initial e and I are uniformly distributed within a range of $0 \leq e \leq 0.07$ and $0 \leq I \leq 0.035$ radians. The distributions follow the relation of $e = 2I$, but the initial values of e and I for each planet are independent. The arguments of pericenter, ω, longitudes of ascending node, Ω, and mean anomalies, M are distributed uniformly in the range $0 \leq (\omega, \Omega, M) < 2\pi$.

Objects whose orbital distance exceeds 100 au are removed from the simulations.

The central bodies of each system have their masses and radii taken from the Kepler data stored in NASA's exoplanet archive. The time-steps used in the simulations are set to be 1/20th of the shortest orbital period of the system. Each simulation is run for 10^7 yr.

4.4. Results

4.4.1. Stability of the inner planetary system templates

Before embarking on a study of how the inner planetary systems are perturbed by the outer planets, we begin by considering the dynamical stability of the inner system templates in the absence of the outer systems. This provides a control set for the simulations, and demonstrates that the instabilities discussed later in this chapter are caused by the outer ice/gas giants. We carried out the stability check for all the selected inner planetary system templates described in section 4.3.1. Adopting the inner eccentricity and inclination distributions mentioned in section 4.3.3 (i.e. $\sigma_e = 0.035$ and $\sigma_I = 0.0175$), the simulations show a number of the Kepler system templates are prone to instability within 10^7 yr.

Figure 4.4 shows the results for the stability test. 100% of the simulations for `Kepler32`, `55`, `238`, and `296` were stable, while $\geq$ 97% of the simulations for `Kepler62`, `84`, `154`, and `186` were stable. The simulations for `Kepler33`, `102`, `169`, and `292` showed higher levels of instability, and for this reason we remove these systems from further consideration. We note that the purpose of this study is not to specifically assess the stability of the Kepler 5-planet systems, since we have made assumptions about the mass-radius relationship and the initial eccentricities and inclinations that may not apply to each of the systems separately. Instead, our aim is to obtain a sample of stable inner systems, given our assumptions, that can then be evolved in the presence of outer systems of giants to examine whether or not the induced perturbations lead to inner systems similar to those that have been observed.

From now on we consider only the eight inner systems: `Kepler32`, `55`, `62`, `84`, `154`, `186`, `238` and `296`. The K-values, that quantify the mutual separations between planets in these systems, are listed in table 4.2, along with the stellar masses. The minimum value in the table is $K_{\rm min} = 10.3$, which is greater than the $K_{\rm min} = 7.1$ required to ensure stability in a 5-planet system with circular orbits for up to 10^6 years (Wu et al., 2019), in agreement with the results of our stability tests.

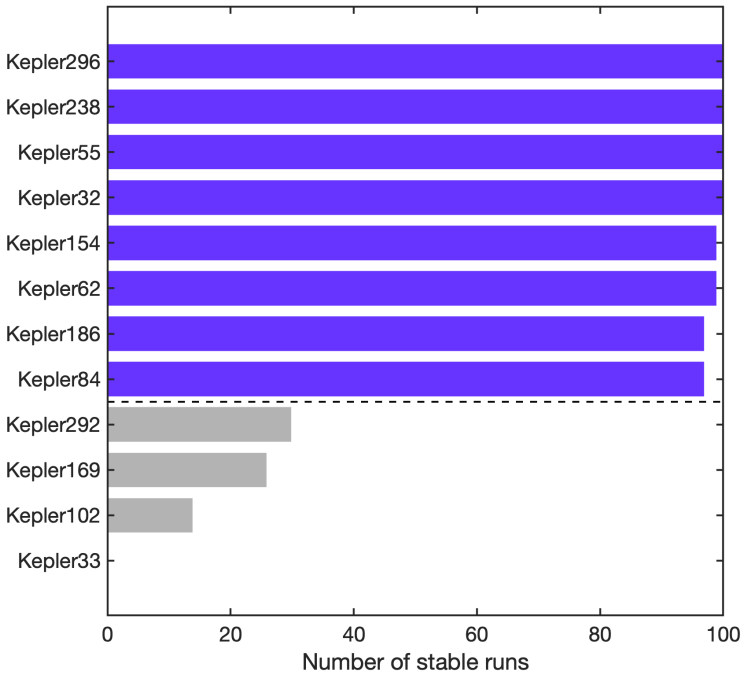

Figure 4.4.: Stability control test for the selected Kepler 5-planet inner system templates. The histogram shows the number of stable runs for each template out a total of 100 simulations. Systems with high levels of stability within 10^7 yr are marked by a purple bar, and the rest are marked in grey.

Table 4.2.: K-values of the stable Kepler 5-planet system templates. K_1 to K_4 denote the K-value from the 1st to 4th pair of adjacent planets, respectively, in order of increasing orbital radius. $\langle K \rangle$ denotes the arithmetic mean of K for the system. Starred values are the minimum K-values in the system, $K_{\rm min}$. The superscripts refer to the following references for the stellar masses: a) Steffen et al. (2013); b) Batalha et al. (2013); c) Xie (2014).

System	K_1	K_2	K_3	K_4	$\langle K \rangle$	$M_\odot$
Kepler-32	50.4	17.5	22.7	13.3*	26.0	0.58
Kepler-55	23.9	25.6	26.7	10.3*	21.6	0.62$^{\rm a}$
Kepler-62	35.4	13.9*	52.8	26.8	32.2	0.69
Kepler-84	24.6	11.6*	21.0	13.9	17.8	1.00
Kepler-154	34.5	17.4	10.9*	16.4	19.8	0.89$^{\rm b}$
Kepler-186	15.8*	28.5	18.0	66.9	32.3	0.54
Kepler-238	33.7	20.3	10.8*	14.9	19.9	1.06$^{\rm c}$
Kepler-296	17.2	16.2	14.8*	18.0	16.5	0.50

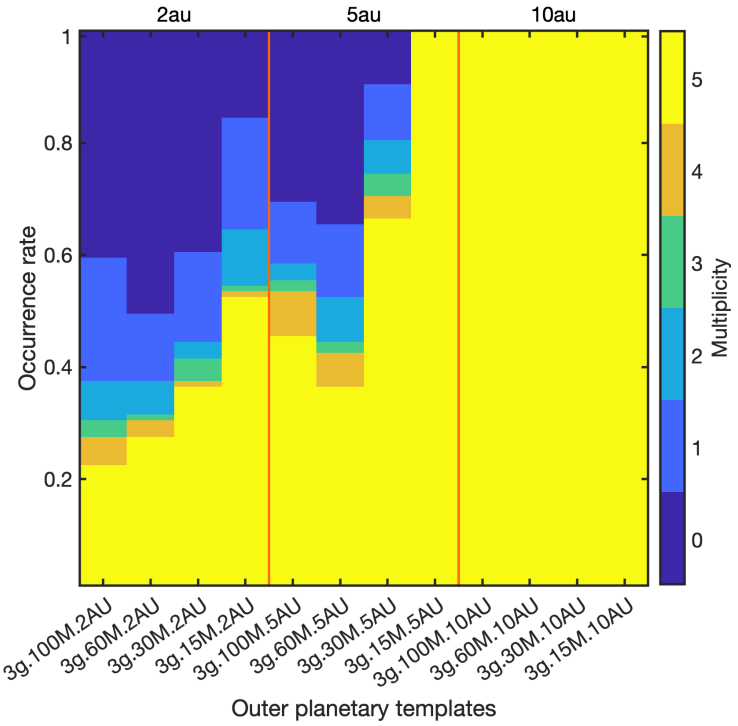

Figure 4.5.: Occurrence rate for the final inner planet multiplicities for $N_{\rm out} = 3$ (3g) runs on `Kepler55`. The left panel is the templates with $\tilde{a}_{\rm out} = 2$ au, middle panel is the templates with $\tilde{a}_{\rm out} = 5$ au, and right panel is for the $\tilde{a}_{\rm out} = 10$ au templates. Different colour represent their relative final multiplicity. Multiplicity of 5 (yellow bar) illustrate unperturbed system and 0 (dark blue bar) for completely destroyed system.

4.4.2. Perturbations by the outer ice/gas giants

It would be time consuming and beyond our computational capabilities to run 100 instances for all combinations of the inner and outer planetary system templates. To optimise the computational resources, we selected one of the eight stable inner templates (`Kepler55`) and ran simulations combining this with all the outer system templates listed in table 4.1. We ran 100 simulations for each configuration. The purpose here is to determine which of the outer planet system configurations give the most promising results, so that we can then focus on these systems in a more comprehensive study.

The final inner system multiplicities from this set of simulations are shown in figures 4.5, 4.6, and 4.7, which show the results from the runs with 3, 6 and 12 outer planets, respectively. In figure 4.5 we can see a drop in the final multiplicities as a function of $\tilde{a}_{\rm out}$. Furthermore, for a given value of $\tilde{a}_{\rm out}$, we can also see a decrease in final multiplicities as the masses of the outer planets increases. As expected, larger

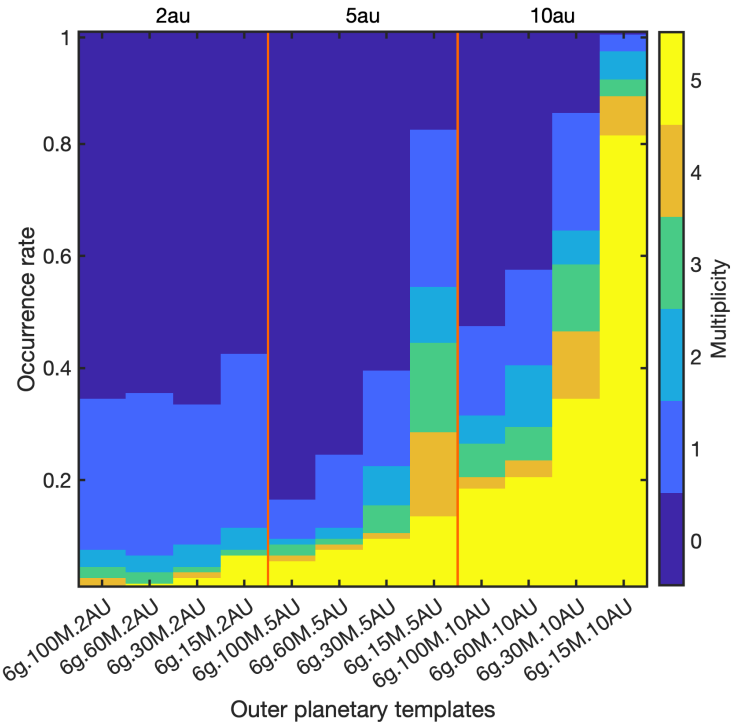

Figure 4.6.: Same as figure 4.5 but for $N_{\rm out} = 6$ (6g) outer templates.

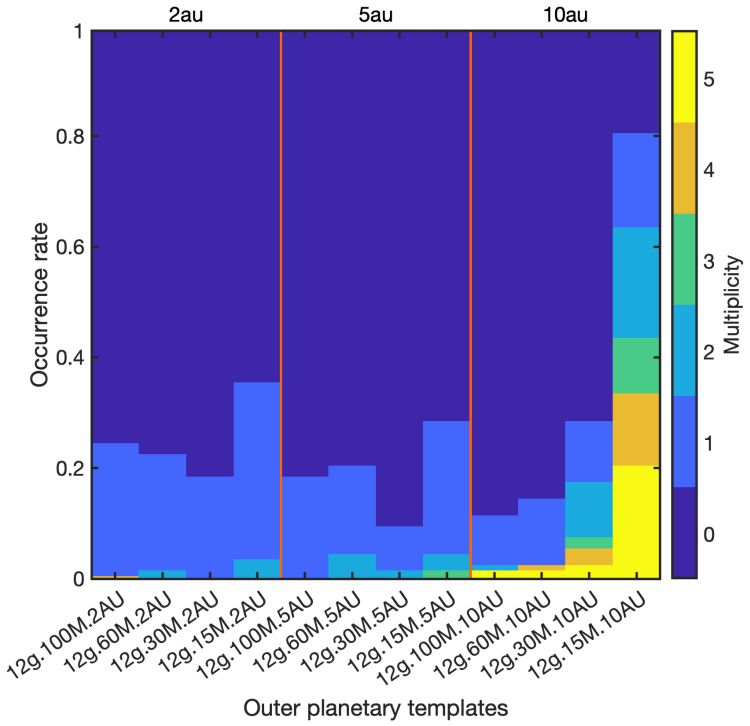

Figure 4.7.: Same as figure 4.5 but for $N_{\rm out} = 12$ (12g) outer templates.

planet masses induce a greater degree of scattering, and closer orbits increase the probability of outer planets having close encounters with the inner planets systems. The same behaviour is also seen in figures 4.6 and 4.7.

For the $N_{\rm out} = 3$ templates (figure 4.5), the subset with $\tilde{a}_{\rm out} = 10$ au shows no perturbation of the final multiplicities of the inner systems at all. Hereafter, we refer to these kinds of systems, where the initial and final values of inner multiplicity are the same ($N_{\rm in,init} = N_{\rm in,final}$), as 'unperturbed systems'. The `3g.15M.5AU` template also produces unperturbed systems only. The remaining outer templates show that the occurrence rate of unperturbed systems is between $\sim 30\%$ to $\sim 75\%$, while the occurrence rate of 'completely destroyed systems' ($N_{\rm in,final} = 0$) is in the range of $< 10\%$ to $\sim 50\%$.

The $N_{\rm out} = 6$ templates (figure 4.6) show a very different occurrence rate of unperturbed and completely destroyed systems compared to the $N_{\rm out} = 3$ templates, and more generally there is a very obvious trend towards greater degrees of perturbation of the inner systems as the number of planets in the outer systems increases. For $N_{\rm out} = 6$, the occurrence rate of completely destroyed systems covers a wide range between ~ 20 to 100 per cent, while the unperturbed systems also shows a wide variety, from 0 to $\sim 80\%$. Meanwhile, the proportion of final multiplicities being equal to 2, 3, and 4 are also higher when comparing to the $N_{\rm out} = 3$ templates. Figure 4.7 shows the final multiplicities for the $N_{\rm out} = 12$ templates. Most of these templates resulted in $\sim 80\%$ of the systems being completely destroyed. The templates with $\tilde{a}_{\rm out} = 10$ au are the only ones that result in unperturbed and moderately perturbed systems among the $N_{\rm out} = 12$ templates.

The results shown in figures 4.5, 4.6, and 4.7 agree with our expectations. Much stronger perturbations are experienced by the inner systems when the outer system planets are: i) closer; ii) more massive; iii) more numerous. The probability of having a close encounter between inner and outer planets is obviously larger for closer in outer systems, and greater degrees of scattering are expected when the outer planets are either more numerous or more massive.

Based on the results shown in figures 4.5, 4.6, and 4.7, we have selected systems with six outer planets and with $\tilde{a}_{\rm out} = 5$ and 10 au for a more in-depth study. As highlighted in table 4.1 and figure 4.3, in total this sample contains eight outer system templates: `6g.15M.5AU`, `6g.30M.5AU`, `6g.60M.5AU`, and `6g.100M.5AU` for $\tilde{a}_{\rm out} = 5$ au, and `6g.15M.10AU`, `6g.30M.10AU`, `6g.60M.10AU`, and `6g.100M.10AU` for $\tilde{a}_{\rm out} = 10$ au. The reason for choosing these systems is that they covered the widest range of final multiplicities in our previously described runs (see figure 4.6). For example, all the runs ofor `6g.100M.5AU` show a perturbed inner system, while

`6g.15M.10AU` shows a majority of unperturbed systems.

4.4.3. Evolution of Kepler templates with 6-planet outer systems

As described in section 4.4.2, the templates consisting of 6 outer planets centred around both 5 and 10 au provide the eight templates we investigate further. Together with the eight Kepler inner system templates described in section 4.4.1, there are 64 systems that we now focus on. For each system we run 100 simulations.

Dynamical evolution

The perturbations arising from the chaotic dynamics of the outer planets can lead to very diverse outcomes for a given combination of inner and outer system templates. Figures 4.8 and 4.9 show the evolution for three different runs from the system template of `Kepler55.6g.60M.10AU`, where the top panel shows the semi-major axis of the planets, the middle panel shows the distance of the pericentre of the planet's orbit to the host star, $r_{\rm peri}$, and the bottom panel shows the evolution of the eccentricities.

Figure 4.8 shows a run where a single planet survives in the inner system. Approximately 1000 years after the start of the simulation, the outer system becomes unstable. This is as expected and compares well to the instability timescales obtained by Chatterjee et al. (2008) when $K = 4$. As the eccentricities of some of the outer planets increase, their pericentre distances decrease and start to approach the inner planets. At $t \approx 4 \times 10^6$ yr one of the outer planets has $r_{\rm peri} < 0.4$ au (thin orange line), strongly perturbing the inner system. Two inner planets are scattered to high eccentricities ($e_{\rm p} \gtrsim 0.8$), resulting in a collision that forms a single planet with $e_{\rm p} \sim 0.3$. The remaining inner planets collide with the giant interloper, an outcome that is common during the simulations. If we were to observe the final state of this system using a transit survey and a viewing position which allows the remaining inner planet to transit its host star, we would classify this system as a being a single, high eccentricity super-Earth. The outer planets would not generally be detectable because of their orbital inclinations with respect to the inner system, and because their orbital periods are too long for multiple transits to be detected within the few years of operation of a Kepler-like survey. Furthermore, the outer planets are below the detection threshold of an RV survey.

The left column of figure 4.9 shows the dynamical evolution of an unperturbed

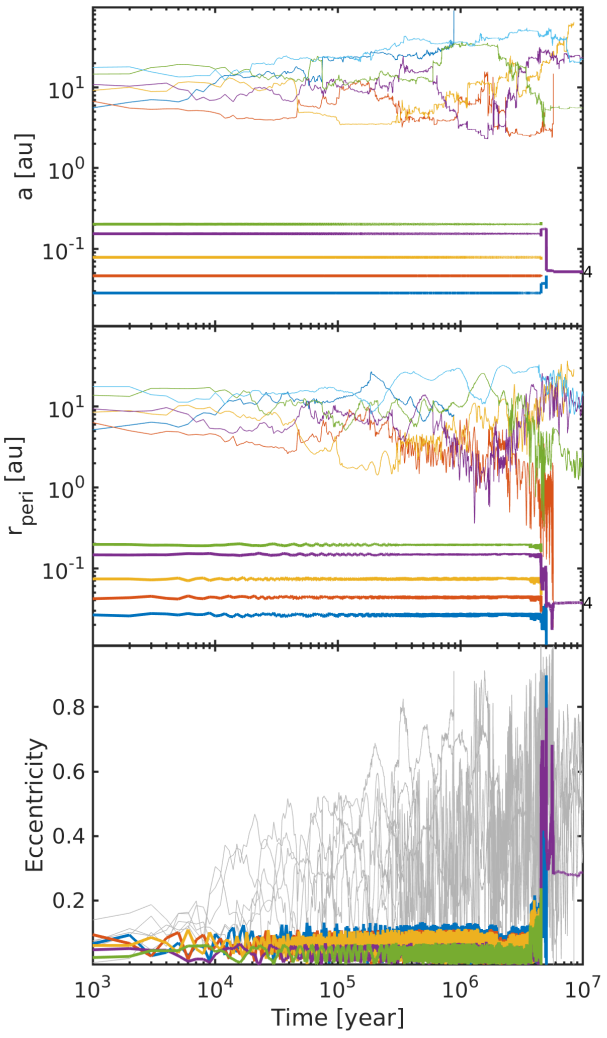

Figure 4.8.: Dynamical evolution during a run for the `Kepler55.6g.60M.10AU` template. Top panel: evolution of planet semi-major axes. Middle panel: evolution of the pericentre distances. Bottom panel: evolution of the eccentricities. Inner system planets are marked in thick solid lines, the numbers labelling the final remaining planets are marked on the right hand side of each plot. Outer system planets are marked using thin solid lines. This evolution show a system in which only one eccentric inner planet survives ($e_{\rm p} \sim 0.3$).

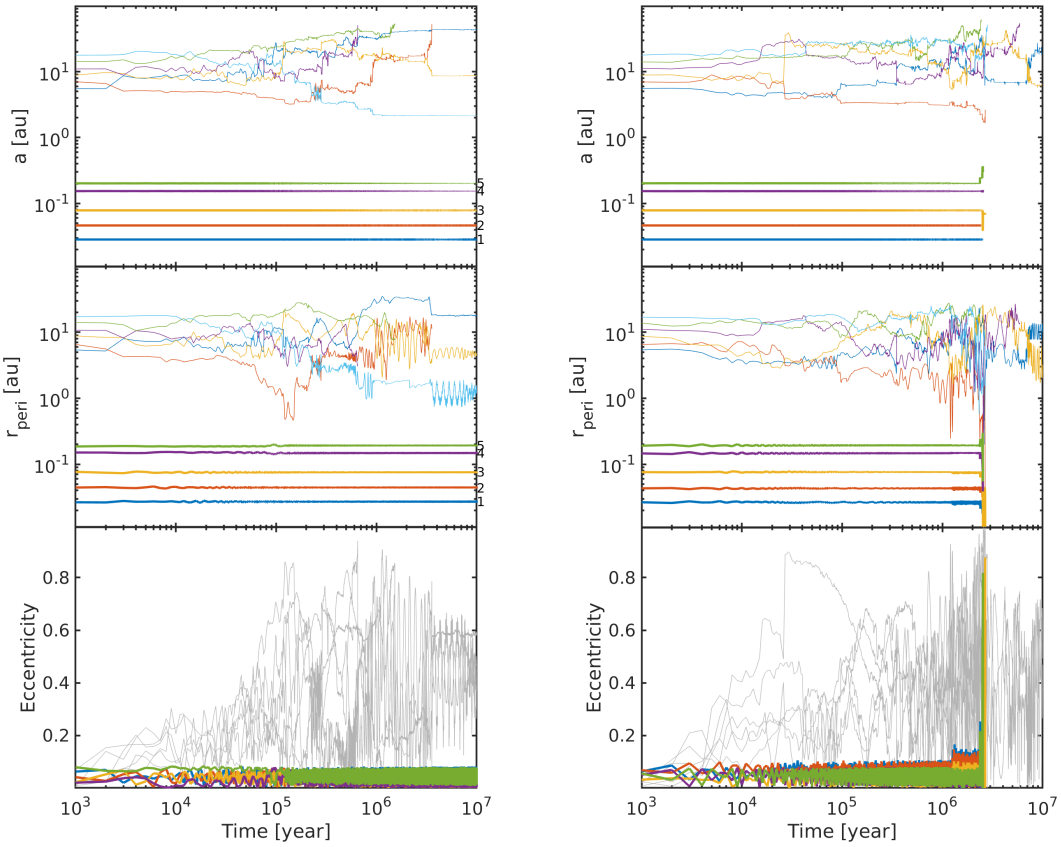

Figure 4.9.: Dynamical evolution during two different runs for the template of `Kepler55.6g.60M.10AU`, similar to figure 4.8. The left panels show an unperturbed system in which all five inner planets survive. The right panels show a completely destroyed system.

system. Even through we label it as unperturbed, the outer planets still experience chaotic evolution due to the small initial Hill separations. The difference compared to the previously described run is simply that the outer planets in this simulation did not make an excursion into the inner system during the chaotic phase. The value of r_{peri} shows that one of the outer planets got as close as ~ 0.5 au (thin orange line), while the outermost inner planet is sitting at ~ 0.2 au (thick green line). The Hill separation of this planet pair was greater than 18 throughout the simulation, leading to only a small perturbation of the inner system, as demonstrated by the evolution of the eccentricities that remain at essentially their initial values. Again, if we were to observe the final state of this system using a transit survey, there would be a finite chance of detecting it as a multi-planet system with low eccentricities, depending on the viewing angle. Together with the high eccentricity single planet system that we discussed above, these two systems demonstrate how different evolutionary

paths originating from similar initial conditions can in principle explain results such as those presented by Xie et al. (2016), for which the mean eccentricity of single-systems is $\langle e_1 \rangle \approx 0.25 - 0.3$ and for multi-systems is $\langle e_{\geq 2} \rangle \approx 0.05$.

An example of evolution leading to a completely destroyed system is shown in the right column of figure 4.9. Similar to the run shown in figure 4.8, the chaotic evolution in the outer system reduces $r_{\rm peri}$ so that an outer planet penetrates into the inner system at $\sim 2.5 \times 10^6$ yr. This induces an instability within the inner system, and allows one of the inner planets (thick yellow line) to accrete the other four inner planets before colliding with the host star.

Overall in our simulations, the dominant mechanism that generates perturbed inner systems is strong scattering of outer giants, leading to one or more giant planet being scattered sufficiently that it passes within the inner system during pericentre passage ($> 95\%$ of the perturbed systems show orbit crossing involving an outer giant and the inner system). Some previous studies have suggested that perturbations of inner system can arise because of secular interactions (e.g. Matsumura et al., 2013), but this mechanism is not commonly seen in our study.

Multiplicities and eccentricities of the inner systems

Before we undertake synthetic observations of the planetary systems resulting from the N-body simulations, and compare them with the observations, we discuss some of their intrinsic properties. We find the multiplicities and eccentricities in the final planetary systems to be very diverse. Figure 4.10 shows the final multiplicities obtained from each of the templates, and these can be compared to the 5 and 10 au subsets in figure 4.6 which apply only to the `Kepler55` system. The overall multiplicities obtained from all 8 inner system templates are very similar to those obtained from the `Kepler55` system, indicating that the dynamics of the outer systems are the main controller of the evolution of the inner systems.

The trend in the multiplicities displayed by the different outer system templates is not unexpected. Planet-planet scattering is more effective for systems containing more massive planets, and planets in the outer systems are more easily able to penetrate into the inner systems if they orbit closer to the central star. Hence, the correlations shown in figure 4.10 between final multiplicities of the inner systems and the properties of the outer systems are easily understood. The `6g.15M.10AU` template shows the largest fraction of unperturbed systems at $\sim 75\%$, and the `6g.100M.5AU` template produced the largest fraction of inner systems that were completely destroyed system, again $\sim 75\%$ of the total. The `6g.15M.5AU` template

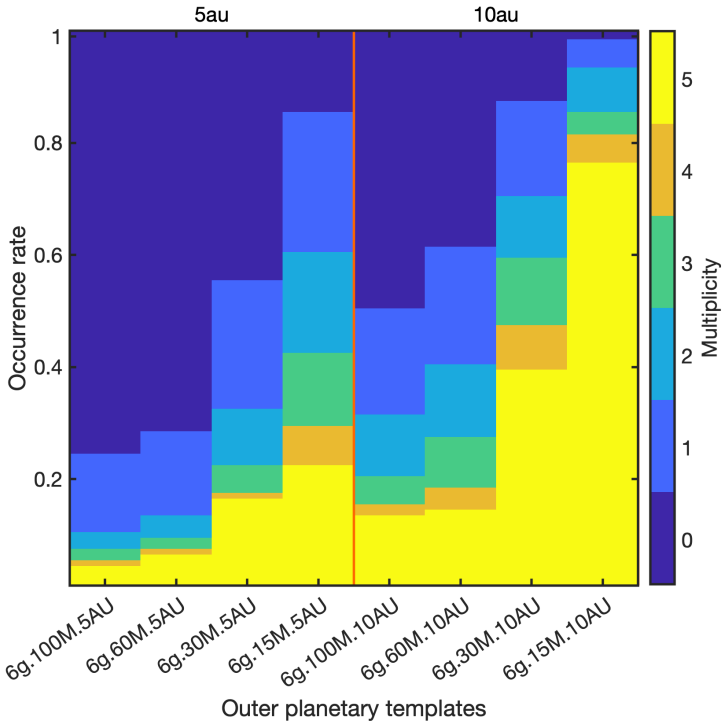

Figure 4.10.: Multiplicity occurrence rates from the simulations of the 8 selected outer system templates combined with all the Kepler templates. The left-most histograms are for systems with $\tilde{a}_{\rm out} = 5$ au, and right-most histograms are for the $\tilde{a}_{\rm out} = 10$ au.

produced the largest fraction of systems that were perturbed but not completely destroyed (i.e. $1 \leq N_{\rm in,final} \leq 4$).

The final multiplicities of the inner systems at the end of the simulations reflect different dynamical histories, and in general one might expect higher multiplicities to arise in systems that have experienced smaller perturbations from the outer planets. One might also therefore expect the final eccentricity distributions to correlate with the multiplicities. Figure 4.11 shows eccentricities versus semi-major axes from all 64 inner/outer template combinations, with each panel showing data for the different multiplicities (one to five planets). Table 4.3 also lists the mean and median eccentricities, $\langle e \rangle$ and $\tilde{e}$, for different multiplicities at the end of the simulations. These values, and figures 4.11 and 4.12, show a clear inverse relation between the final multiplicities and the values of $\langle e \rangle$ or $\tilde{e}$, as expected. Figure 4.12 shows the final eccentricity distribution for the unperturbed systems (5-planets, red solid line) has remained similar to the initial eccentricity distribution (Rayleigh distribution with $\sigma_e = 0.035$, red dashed line). This can also be compared to the eccentricity distribution that arose from the simulations performed to check the stability of the

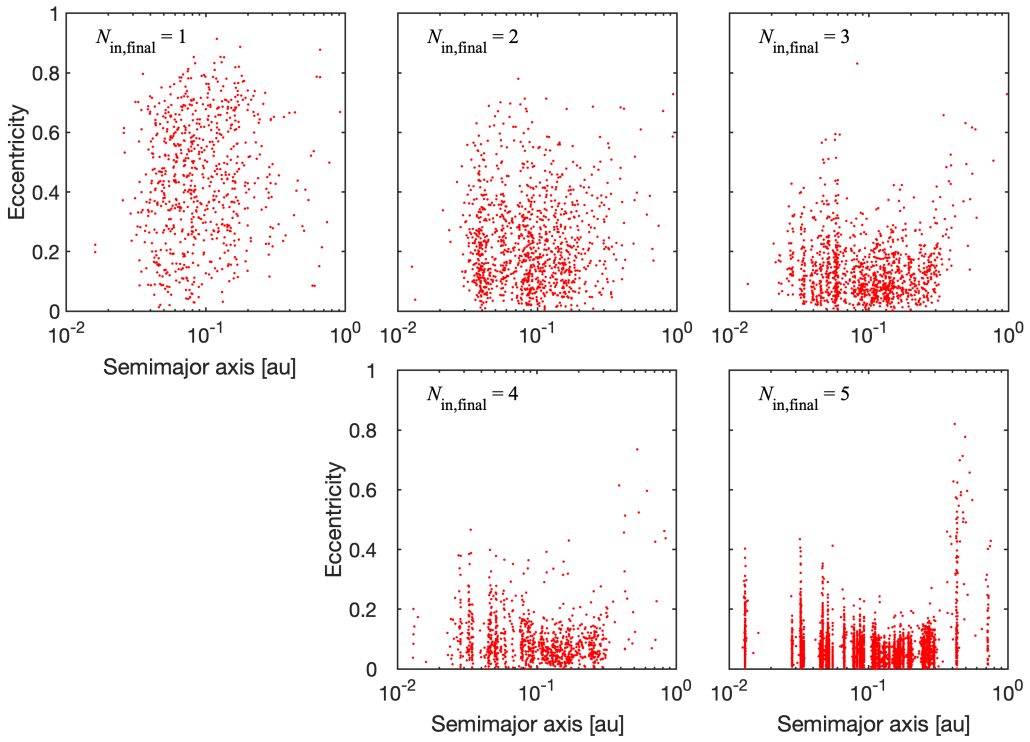

Figure 4.11.: Scatter plots of e as a function of a for the final inner planetary systems. Different panel shows the distribution of different final inner system multiplicities (from the top-left panel for $N_{\rm in,final} = 1$ in order to bottom-right panel for $N_{\rm in,final} = 5$).

Kepler templates, described in section 4.4.1 (denoted by the grey dashed line in figure 4.12). The similarity between these three distribution shows that if the inner system retains the original multiplicity, the perturbations from the outer systems are small and do not significantly excite the inner systems.

The single planet systems (blue solid line in figure 4.12 and left panel in figure 4.11) are the most eccentric on average. The e-distribution for 1-planet systems is well fitted by a Rayleigh distribution with $\sigma_e = 0.410$ for $e < 0.6$ (figure 4.12, magenta dashed line), but these systems did not provide enough high eccentricity planets with $0.6 < e < 1$ to fit the large eccentricity end of the Rayleigh distribution. The e-distribution for 2-planet systems (green solid line) is similar to the Rayleigh distribution with $\sigma_e = 0.167$ (blue dashed line, the suggested Rayleigh distribution eccentricity parameter for observed single-planet system suggested by Mills et al. (2019)). We note that we are looking at the intrinsic properties of the systems here, and not those derived from a set of synthetic transit observations, so the fact that the simulated 1-planet systems do not match the Rayleigh distribution for $\sigma_e = 0.167$ is not particularly relevant, as we discuss later in this chapter when we examine the

Table 4.3.: Mean ($\langle e \rangle$), median ($\tilde{e}$), and standard deviations (σ) of the eccentricity of planets in the final inner systems from the focused outer templates runs. Different column represent different final inner multiplicities. The eccentricity distributions with values of $\langle e \rangle$, $\tilde{e}$, and σ listed here are shown in figures 4.11 and 4.12.

	Multiplicity				
	1	2	3	4	5
$\langle e \rangle$	0.48	0.24	0.15	0.10	0.06
$\tilde{e}$	0.48	0.21	0.12	0.08	0.05
σ	0.21	0.16	0.11	0.09	0.06

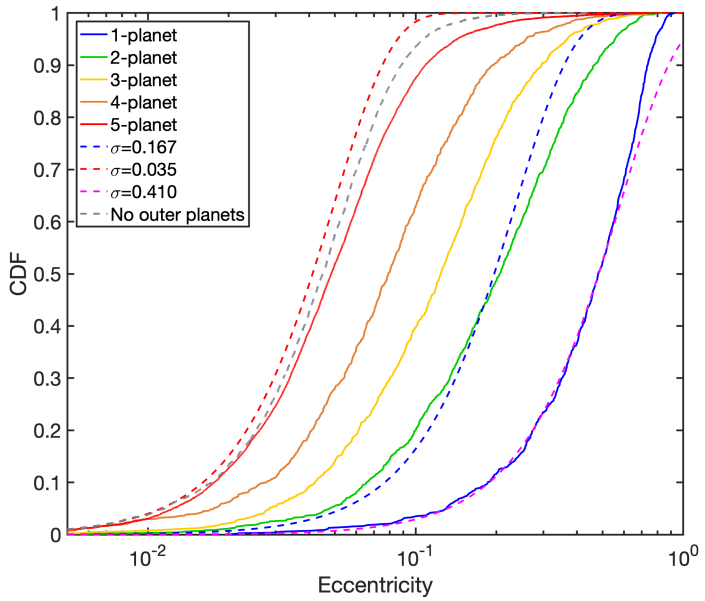

Figure 4.12.: Cumulative distribution functions, CDFs, of the final inner system planet eccentricities obtained from the focused templates simulations. The solid lines correspond to the distribution of different $N_{\rm in,final}$-planet systems. For comparison, the blue, red, and purple dashed line show the distribution of eccentricities drawn from Rayleigh distributions with eccentricity parameters $\sigma = 0.167$, 0.035, and 0.410 respectively. The gray dashed line is the CDF drawn from the set of control (with no outer planetary system, section 4.4.1).

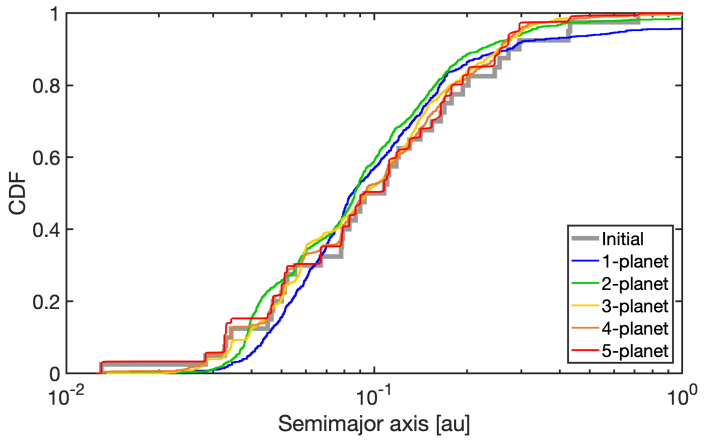

Figure 4.13.: CDFs of the inner planet semi-major axes obtained from 64 different combinations of inner/outer templates. Different line colours represent different multiplicities. The grey line represents the CDF of the initial semi-major axes.

results of synthetic transit observations of the simulated systems (see section 4.5).

Figure 4.13 shows the cumulative distribution functions for the final semi-major axes of the simulated planets as a function of the final multiplicity. Unlike in figure 4.12, the semi-major axis distributions do not vary strongly with multiplicity, and are very similar to the initial values. The right panel of figure 4.11 shows a pattern of vertical strips indicating that the planets in the unperturbed systems did not move away from their original semi-major axes significantly. Furthermore, we also observe a tendency for the innermost and outermost planets to be the most eccentric for the systems containing 3, 4 and 5 planets. The outermost planets in these systems have experienced the strongest perturbations due to the outer planets, and hence show enhanced eccentricities. The innermost planets that display the largest eccentricities, however, obtained these larger values because of the redistribution of angular momentum deficit (AMD, Laskar and Petit, 2017) within some of the inner planetary systems during the evolution described in Sect. 4.4.1, prior to the runs being performed with the giant planets having been inserted. This point is discussed further in appendix A.4. For lower multiplicities these patterns becoming increasingly indistinct, which together with the eccentricity distributions mentioned above, indicates that the planets in small multiplicity systems have experienced stronger scattering, as expected.

Table 4.4.: Final multiplicity occurrence rates for a selection of outer system templates (where planets have final $a > 1$). Rates in the right-most column are the sum of the occurrence rates for multiplicities of 2 and 3. Note that systems recorded as having zero giant planets actually have surviving planets with $a < 1$.

	Multiplicity occurrence rate (%)							
Outer template	0	1	2	3	4	5	6	{2,3}
`3g.15M.2AU`	0	4	82	14	-	-	-	96
`3g.30M.2AU`	0	16	82	2	-	-	-	84
`3g.60M.2AU`	2	56	42	0	-	-	-	42
`3g.100M.2AU`	1	64	54	1	-	-	-	55
`6g.15M.2AU`	0	7	30	44	19	0	0	74
`6g.30M.2AU`	0	18	58	24	0	0	0	82
`6g.60M.2AU`	0	47	49	4	0	0	0	53
`6g.100M.2AU`	1	64	34	1	0	0	0	35
`6g.15M.5AU`	0	0	0	2	25	56	17	2
`6g.30M.5AU`	0	2	14	37	44	3	0	51
`6g.60M.5AU`	0	16	53	30	1	0	0	83
`6g.100M.5AU`	0	39	53	8	0	0	0	61
`6g.15M.10AU`	0	0	0	2	3	38	57	2
`6g.30M.10AU`	0	0	1	21	37	37	4	22
`6g.60M.10AU`	0	12	35	40	12	1	0	75
`6g.100M.10AU`	0	27	61	11	1	0	0	72

4.4.4. Multiplicities and eccentricities of the outer systems

The outer planet templates we consider in this study are at or below the RV detection limit (assumed to be 3 ms^{-1}) and are not amenable to transit surveys. Hence, the outcomes of our N-body simulations cannot be compared with observations. Previous studies have focused on outer giant systems comprised of planets that are above the RV detection limit (e.g., Chatterjee et al., 2008; Petrovich et al., 2014; Huang et al., 2017). Jurić and Tremaine (2008) suggested that the dynamical evolution of initially high multiplicity giant systems ($N_{\rm init} > 3$ and $m_{\rm p} > 0.1$ $M_J \sim 32$ $M_\oplus$) are likely to result in a lower multiplicity ($N_{\rm final} = \{2, 3\}$). A subset of our simulations with initial conditions lying within those ranges (e.g., `6g.60M.5AU` and `6g.100M.10AU`) agree with Jurić and Tremaine (2008), where the final systems with $N_{\rm out,final} = \{2, 3\}$ make up the majority (see right column of table 4.4). On the other hand, a subset of our simulations do not lie within the parameter space studied by Jurić and Tremaine (2008). Our outer system templates include planets with masses smaller than 0.1 M_J. For those systems (e.g. `6g.15M.5AU` and `6g.30M.10AU`), the majority of systems no longer end up with multiplicities of 2 or 3. For example,

the template `6g.15M.5AU` ended up with only 2 % of systems having 2 or 3 planets. A correlation between the final multiplicities and the masses of the planets can be seen in table 4.4, where the lower mass planets maintain a higher final multiplicity due to less efficient scattering.

The upper panel of figure 4.14 shows the eccentricities versus semi-major axes of the outer giant planets from the `Kepler55.3g` ($N_{\rm out,init} = 3$) runs. According to table 4.4, most of the systems in this subset end up with $N_{\rm out,final} = 2$. This agrees with previous studies that considered three outer planets, where $N_{\rm out,final} = 2$ is the most common outcome (e.g., Petrovich et al., 2014; Huang et al., 2017). Furthermore, the 2-planet systems show a v-shaped distribution, which can also be seen in the simulation results of Chatterjee et al. (2008), Petrovich et al. (2014), and Huang et al. (2017). Similar to Chatterjee et al. (2008), we demonstrate that there exists a separation between the inner and outer planets in the population of the 2-planet systems. A clear division between the two sub populations can be seen, where the inner planet tends to follow the apoapsis, $a = a_{\rm in}/(1+e)$, of the initial innermost planet (left dashed line in figure 4.14), and the outer planet tends to follow the periapsis, $a = a_{\rm Mid}/(1-e)$, of the middle planet of the three original planets (right dashed line). The $a_{\rm in}$ and $a_{\rm Mid}$ that applied to the dashed lines in figure 4.14 are 1.51 and 2.00 au respectively, where this value came from the semi-major axes of the innermost and middle planets in the outer systems that make up the `Kepler55.3g.100M.2AU` template.

The v-shaped distribution seen in the upper panel cannot be seen in the results from the higher initial multiplicity templates, such as `Kepler55.6g.2AU` (figure 4.14, bottom panel), because the larger amount of scattering washes this feature out. This feature can also not be seen in the e versus a plot for the observed RV planets shown in figure 4.15, indicating that this population is not consistent with it being the result of scattering from initial conditions similar to those in the `Kepler55.3g` ($N_{\rm out,init} = 3$) runs.

4.5. Synthetic observation of the final planetary systems

The multiplicity of a planetary system observed using the transit method depends on the viewing angle, intrinsic multiplicity and mutual inclinations between the planets within the system, the radius of the host star and orbital radii of the planets (where we have ignored the effect of the finite planet radius). To compare the outcomes

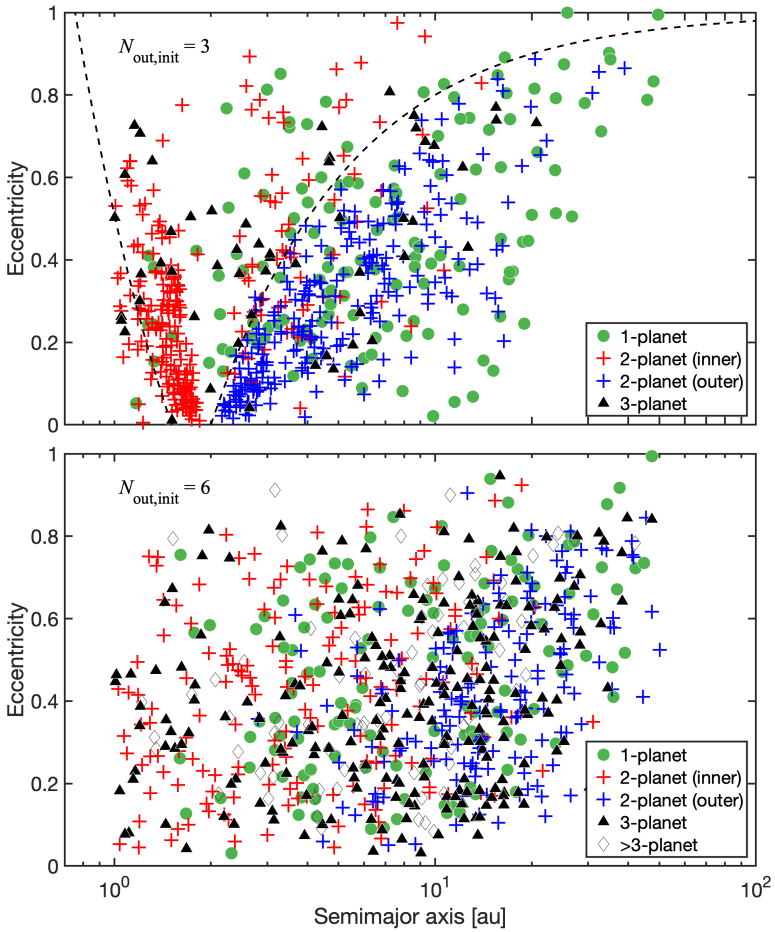

Figure 4.14.: Scatter plots e versus a for the final outer planets resulting from a subset of the `Kepler55` runs. Different symbols represent the different values of $N_{\rm out,final}$. For the $N_{\rm out,final} = 2$ case, the inner planets are marked in red while the outer planets are marked in blue. The upper panel shows results for systems where the outer systems had $N_{\rm out,init} = 3$ planets and $\tilde{a}_{\rm out,init} = 2$ au. The left dashed line denotes the value of $a = 1.51/(1 + e)$ and the right dashed line denotes the value $a = 2.0/(1 - e)$. The bottom panel shows the e versus a distribution from a subset of outer templates with $N_{\rm out,init} = 6$ and $\tilde{a}_{\rm out,init} = 2$ au.

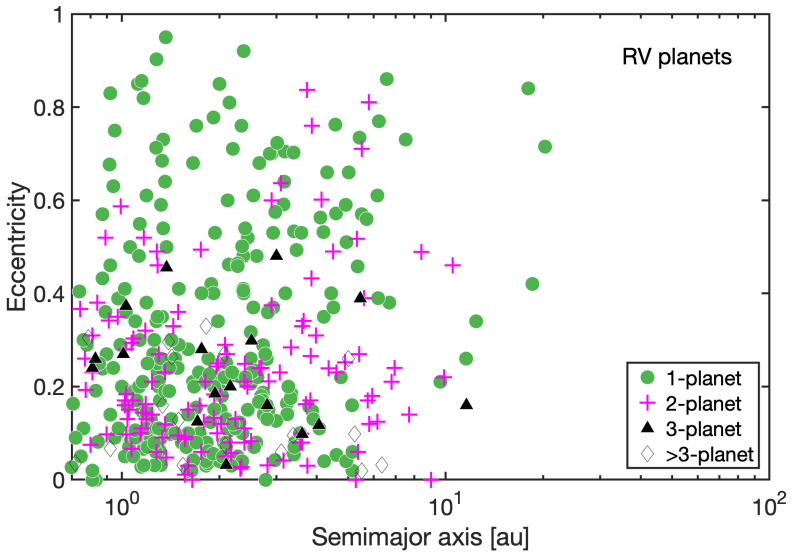

Figure 4.15.: Eccentricities versus semi-major axes of the observed RV planets. 1-planet systems are marked as green circles, 2-planet systems are marked as pink crosses, 3-planet systems are marked as black triangles, and RV systems with multiplicities higher than three are marked with unfilled black diamonds.

of our N-body simulations with analyses of Kepler data relating to the multiplicity and eccentricity dichotomies, we now present the results of synthetically observing the simulation results.

Each simulated planetary system is synthetically observed from 10,000 randomly chosen viewing locations, isotropically distributed with respect to each host star. We only consider planets that are within the inner planetary system ($a_{\rm p} < 1$ au). As the smallest planet we considered in the simulations (Kepler-62c) is a confirmed Kepler planet, we assume all planets satisfy the observation limits of a Kepler-like survey.

4.5.1. Observed multiplicities

Following the approach of Johansen et al. (2012), we consider the relative numbers of one-planet, two-planet, ..., five-planet systems detected when the simulation outcomes are synthetically observed. Similar to section 1.3.2 and Poon et al. (2020), using the observed numbers of one-planet, two-planet, etc. systems, we then define a Transit Multiplicity Ratio (abbreviated to TMR hereafter) as follows:

$$\mathrm{TMR}(i:j) = \frac{\text{Number of } i\text{-planet systems}}{\text{Number of } j\text{-planet systems}}, \tag{4.5}$$

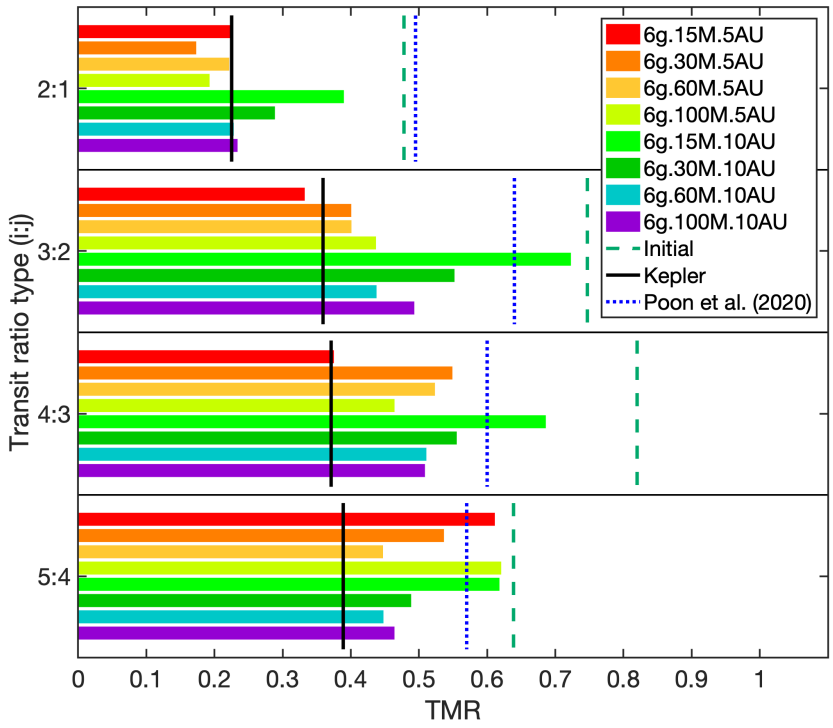

Figure 4.16.: Synthetic TMRs from the simulations. Different coloured bars denote TMRs of different outer templates. The green dashed vertical lines show the TMRs of the initial inner systems, the black vertical lines show the observed Kepler TMRs, and the blue dotted lines show the TMRs obtained from N-body simulations of *in situ* formation of super-Earths in chapter 3.

where i and j represent the numbers of planets detected during each synthetic observation of each system.

The TMR values obtained are shown in figure 4.16. The coloured histograms show the values obtained for each of the outer system templates, the jade vertical dashed lines mark the TMR values of the initial conditions, the black vertical lines mark the TMRs from the Kepler data, and the blue vertical dotted line shows the results obtained in chapter 3 from its simulations of *in situ* formation of super-Earth systems.

The TMRs drop below the initial values for all outer templates and for all multiplicity ratios. For the 2:1 and 3:2 TMRs, the systems with $\tilde{a}_{\rm out} = 5$ au provide somewhat better fits to the Kepler TMRs than those with $\tilde{a}_{\rm out} = 10$ au. Considering the 2:1 TMR panel in particular, the `6g.15M.5AU`, `6g.60M.5AU`, and `6g.60M.10AU` templates have an almost exact match to the Kepler 2:1 TMR. Very close agreement is obtained between the Kepler 3:2 TMR and those produced by the templates

6g.15M.5AU, 6g.30M.5AU and 6g.60M.5AU. Apart from the 6g.15M.5AU template, however, the 3:2 TMRs obtained from the synthetic observations are higher than obtained from the Kepler data, indicating that the simulations are over-producing systems of 3 transiting planets relative to 2 transiting planets. Indeed, for higher multiplicity ratios than 2:1 more generally, the simulated TMRs are higher than the corresponding Kepler TMRs, suggesting that the mutual inclinations within the simulated systems are smaller than among the Kepler systems.

When considering the TMRs across all multiplicity ratios, we see that the template 6g.15M.10AU is consistently a factor of ∼1.5-2 above the corresponding Kepler TMRs, and can be considered the worst performing template. The TMRs for this template remain closer to the initial values compared to the others, because the majority of final systems for this template are unperturbed systems (see figure 4.10), so it is expected that the change in TMRs would be limited. While no template fits the Kepler TMRs for all multiplicity ratios, the best performing overall are 6g.15M.5AU, 6g.60M.5AU, 6g.100M.5AU and 6g.60M.10AU. Each of these provides a decent fit to the Kepler TMRs for three of the multiplicity ratios considered in figure 4.16, and each performs relatively poorly in one of multiplicity ratios.

In figure 4.16, comparison can also be made between the simulated TMRs to the TMRs from the *in situ* formation simulations in chapter 3 and Poon et al. (2020), where the mutual inclinations of the super-Earths are self-excited by gravitational scattering between the planets as they form. In general (except for the 3:2 TMR for 6g.15M.10AU), the TMRs obtained in this current study are in significantly closer agreement with the Kepler observations than those obtained for systems that do not experience perturbations from outer planets.

4.5.2. Eccentricity distributions

In a recent study, Xie et al. (2016) suggested that the mean eccentricity of Kepler single planet systems is significantly higher ($\langle e \rangle \approx 0.3$) than that of the multi-planet systems ($\langle e \rangle \approx 0.04$). More recently, Mills et al. (2019) also showed there is an eccentricity dichotomy, and they obtained $\langle e \rangle \approx 0.21$ for singles and $\langle e \rangle \approx 0.05$ for multi-planet systems.

Recent studies that consider the dynamics of planets in compact inner systems show that a small dichotomy arises between the eccentricities of the single and multi-transiting systems, but these fail to reproduce the large eccentricity values for the 1-planet systems (e.g. chapter 3; Poon et al., 2020; MacDonald et al., 2020). An increase in the orbital eccentricity of a planet can be induced by planet-planet

scattering, but this is limited to producing values $e_{\rm p} \sim v_{\rm e}/v_{\rm K}$, where $v_{\rm e}$ and $v_{\rm K}$ represent the escape velocity from a planet and the Keplerian velocity. Van Eylen et al. (2019) and Poon et al. (2020) argued that higher mass components are needed in order for scattering to produce systems with high enough eccentricities to match the observations (see also chapter 3).

Table 4.5 lists the mean and median eccentricities from our synthetic transit observations. Each template shows a clear eccentricity dichotomy, where the values of the means and medians of the eccentricities for single-transiting systems are always larger than the values for multi-transiting systems ($\{\langle e_1\rangle, \tilde{e}_1\} > \{\langle e_{\geq 2}\rangle, \tilde{e}_{\geq 2}\}$). The dichotomy signal becomes stronger as the masses of the outer planets increases and the initial orbital radii decrease. For example, the `6g.100M.5AU` template gives $\{\langle e_1\rangle, \langle e_{\geq 2}\rangle\} = \{0.25, 0.08\}$, while the template with lower mass, `6g.15M.5AU`, gives $\{\langle e_1\rangle, \langle e_{\geq 2}\rangle\} = \{0.15, 0.07\}$. The template `6g.15M.10AU` has the lowest mass and the most distant outer planets, and only displays a small eccentricity dichotomy signal ($\{\langle e_1\rangle, \langle e_{\geq 2}\rangle\} = \{0.11, 0.06\}$). The value of $\{\langle e_1\rangle, \langle e_{\geq 2}\rangle\}$ for the template `6g.15M.10AU` is only slightly higher than the value for the control set of simulations (only the inner Kepler templates and no outer systems), as the multiplicities shown in figure 4.10 illustrate that ~ 75 % of the runs using this template resulted in an unperturbed system.

Rather than just considering the mean and median values of the eccentricity distributions, Mills et al. (2019) supposed the distributions follow a Rayleigh distribution and concluded that under this assumption the eccentricity parameters $\sigma_e = 0.167$ and 0.035 gave the best fits to the Kepler single and multiple systems, respectively. Figure 4.17 shows the eccentricity distributions of the synthetically observed single and multiple systems (thick solid blue and red lines) from the `6g.60M.5AU` template, together with the Rayleigh distributions suggested for the Kepler single and multiple systems (dashed blue and red lines) (Mills et al., 2019). Similar plots for all eight outer system templates are shown in figures A.14 and A.15 in the appendix, and although they differ in detail the plots show similar behaviours and trends. In figure 4.17 we see that the synthetically observed e-distribution for the single-transiting systems does not follow a Rayleigh distribution very closely, but nonetheless it has a median eccentricity very similar to that of the suggested Rayleigh distribution with $\sigma_e = 0.167$. For the multiple systems, their distributions are closer to a Rayleigh distribution, and we see that as the multiplicity decreases there is a tendency for the eccentricities to increase. For example, the 2-planet systems produce a more eccentric distribution than the 4- or 5-planet systems, and this is because on average the lower multiplicity systems come from observations of systems that have been more

Table 4.5.: Mean (upper table) and median (lower table) eccentricity of the inner systems obtained by the synthetic transit observations in different outer templates. The observed mean eccentricities are listed by their observed multiplicity. The subscript 1 represents the single-transit systems and subscript ≥ 2 represents the multi-transiting systems. The control set did not include outer systems in the runs (see section 4.4.1).

	Mean eccentricity					
Template	$\langle e_1 \rangle$	$\langle e_2 \rangle$	$\langle e_3 \rangle$	$\langle e_4 \rangle$	$\langle e_5 \rangle$	$\langle e_{\geq 2} \rangle$
6g.15M.5AU	0.21	0.10	0.07	0.05	0.06	0.09
6g.30M.5AU	0.23	0.09	0.05	0.05	0.06	0.09
6g.60M.5AU	0.25	0.09	0.06	0.05	0.05	0.07
6g.100M.5AU	0.25	0.08	0.06	0.04	0.05	0.08
6g.15M.10AU	0.11	0.06	0.05	0.04	0.05	0.06
6g.30M.10AU	0.17	0.08	0.06	0.05	0.05	0.08
6g.60M.10AU	0.23	0.11	0.07	0.05	0.06	0.10
6g.100M.10AU	0.22	0.11	0.06	0.05	0.04	0.08
Control	0.08	0.05	0.04	0.04	0.04	0.05

	Median eccentricity					
Template	$\tilde{e}_1$	$\tilde{e}_2$	$\tilde{e}_3$	$\tilde{e}_4$	$\tilde{e}_5$	$\tilde{e}_{\geq 2}$
6g.15M.5AU	0.15	0.08	0.06	0.04	0.06	0.07
6g.30M.5AU	0.17	0.07	0.04	0.04	0.06	0.07
6g.60M.5AU	0.20	0.07	0.05	0.04	0.04	0.06
6g.100M.5AU	0.23	0.07	0.05	0.04	0.05	0.06
6g.15M.10AU	0.08	0.05	0.04	0.04	0.04	0.05
6g.30M.10AU	0.11	0.06	0.05	0.04	0.05	0.06
6g.60M.10AU	0.16	0.09	0.06	0.05	0.06	0.07
6g.100M.10AU	0.16	0.07	0.05	0.04	0.04	0.06
Control	0.07	0.05	0.04	0.04	0.04	0.05

strongly perturbed by the outer planets. Hence, the synthetically observed systems show the same inverse relation between eccentricity and observed multiplicity that was noted in section 4.4.3 when discussing the intrinsic properties of the simulated planetary systems, albeit at a lower level of significance.

Figure 4.17 also shows that the 2-planet systems (yellow line) provide the main contribution to the eccentricity distribution for all multi-planet systems (thick red line), and the effect of this is to shift the eccentricity distribution away from that displayed by the control set that characterises the dynamics of the unperturbed inner systems, to one corresponding to systems that are more dynamically excited. This effect reduces the dichotomy signal between the single and multiple systems that comes out of the synthetic observations because the detected 2-planet systems are generally more excited than the higher multiplicity systems.

Considering both the TMRs and the mean eccentricities together, the template

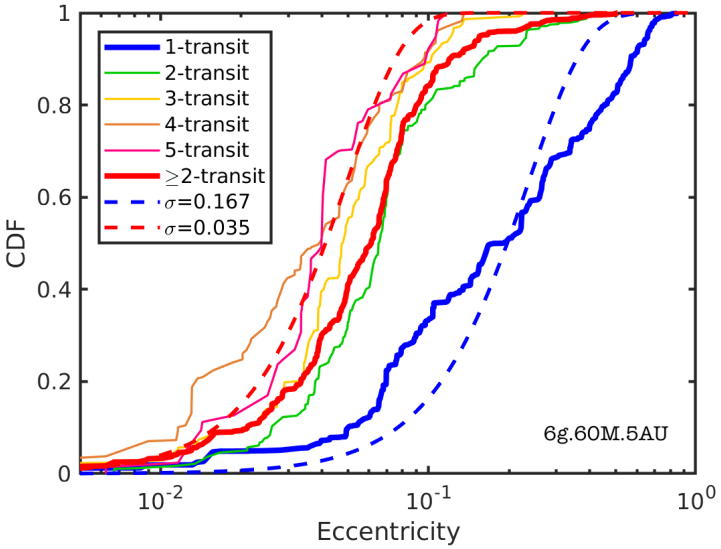

Figure 4.17.: CDFs of the eccentricities obtained from the synthetic observations of the simulations for the `6g.60M.5AU` template. For comparison, the CDFs of eccentricities drawn from Rayleigh distributions with eccentricity parameters $\sigma = 0.167$ and $\sigma = 0.035$ are also plotted.

`6g.60M.5AU` appears to be the best performing overall. The mean eccentricities for single and multiple systems show a strong dichotomy $\{\langle e_1 \rangle, \langle e_{\geq 2} \rangle\} = \{0.25, 0.07\}$ that is in decent agreement with the Kepler eccentricity dichotomy $\{\langle e_1 \rangle, \langle e_{\geq 2} \rangle\} = \{0.3, 0.04\}$ (Xie et al., 2016) or $\{\langle e_1 \rangle, \langle e_{\geq 2} \rangle\} = \{0.21, 0.05\}$ (Mills et al., 2019). The 2:1, 3:2 and 5:4 TMRs are in very good agreement with the Kepler TMRs, with only the 4:3 TMR being in significant disagreement. The template `6g.100M.5AU` also performs well, while for the more distant systems of outer planets centred around 10 au the template `6g.60M.10AU` is the best performing.

4.6. Impact of additional physics

In this section we consider the impact of different physical processes that were not included in the main suite of simulations presented in earlier sections.

4.6.1. Relativistic precession

Precession due to General Relativity (GR) can become a significant effect for planets orbiting close to their host stars, and in particular can influence the secular interactions within planetary systems, as shown for example in the recent study by Marzari and Nagasawa (2020). As discussed in section 4.4.3, the dominant effect

that perturbs the inner systems in our study is the mutual scattering of outer giants, such that during pericentre passage they enter the inner system and cause strong scattering of the planets there. We do not expect GR to provide a stabilising effect in this situation, and this expectation appears to be confirmed by Huang et al. (2017), who examined the influence of GR in their study of planetary scattering.

4.6.2. Planet-planet collisions

Our simulations adopted a perfect merger treatment of planet-planet collisions, and a more realistic prescription might lead to removal of mass from the colliding bodies (Leinhardt and Stewart, 2012; Stewart and Leinhardt, 2012), especially for high-velocity close-in and eccentric collisions. To investigate the changes to our main results that arise when using a more realistic collision model, we have rerun a subset of simulation using a version of SYMBA (Duncan et al., 1998; Poon et al., 2020; Scora et al., 2020), which implements the collision algorithm from Leinhardt and Stewart (2012). Here, the outcome of a collision falls into one of nine regimes: supercatastrophic disruption, catastrophic disruption, erosion, partial accretion, hit-and-spray, hit-and-run, bouncing collision, graze-and-merge, and perfect merger, depending on the collision conditions. The total mass before and after a collision is conserved and obeys the relation of,

$$M_{\rm Total} = M_{\rm LR} + M_{\rm SLR} + M_{\rm Total,debris}, \tag{4.6}$$

where $M_{\rm Total}$ is the total mass of the colliding bodies before the collision, $M_{\rm LR}$ is the mass of the post-collision largest remnant, $M_{\rm SLR}$ is the mass of the second largest remnant, and $M_{\rm Total,debris}$ is the total mass of the debris particles that generated during the collision. For a detailed description of the adopted version of SYMBA, we refer the reader to the model descriptions in section 2.1, Poon et al. (2020) and Scora et al. (2020).

We consider the runs from the subset labelled `6g.60M.5AU` (8 inner planetary systems $\times$ 100 runs each = 800 runs in total). Given we are using SYMBA and not MERCURY for these simulations, we ran the same 800 set of initial conditions using both perfect accretion and the more realistic collision model. Figure 4.18 shows the cumulative distribution function of eccentricities obtained from the synthetic transit observation of the final systems. The different collision models provide similar distributions of the observed eccentricities for both single- and multi-transit systems. They lead to similar mean and median values, where the perfect accretion models

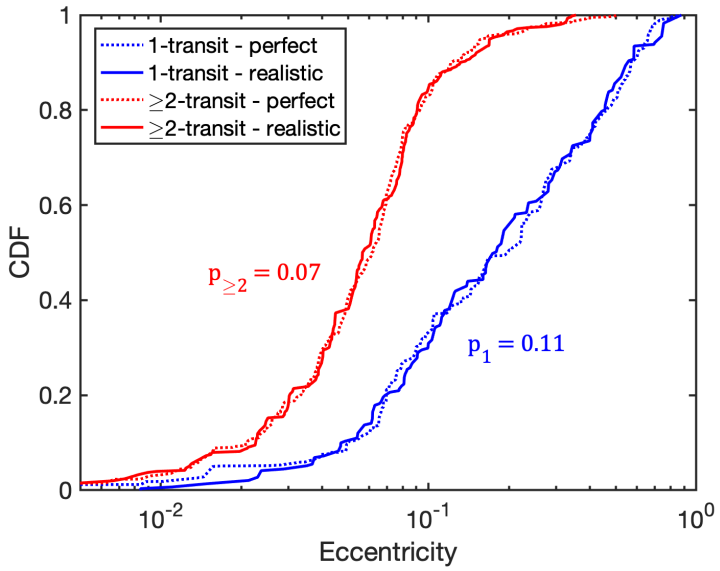

Figure 4.18.: CDFs of the eccentricities obtained from synthetic observations of the simulation outcomes for the `6g.60M.5AU` template by the perfect (dotted lines) and the realistic (solid lines) accretion model. The dotted lines are the same as the thick solid lines from figure 4.17 in the same colours. The p-values from the K-S test are $p_1 = 0.11$ and $p_{\geq 2} = 0.07$.

have $\{\langle e_1 \rangle, \tilde{e}_1, \langle e_{\geq 2} \rangle, \tilde{e}_{\geq 2}\} = \{0.25, 0.20, 0.07, 0.06\}$ (table 4.5) and realistic collision models have $\{0.25, 0.18, 0.07, 0.06\}$. Moreover, the Kolmogorov-Smirnov test (K-S test) between the two sets of synthetically observed eccentricities yields p-values for the single- and multi-transit of $p_1 = 0.11$ and $p_{\geq 2} = 0.07$, respectively. Poon et al. (2020, see also appendix A.1) demonstrates that the re-accretion of the collision debris occurs quickly (on a time-scale of 10^3 yr) and most of the debris (>80 %) is re-accreted back by the largest/second largest remnant. This test shows that our results are insensitive to the collision prescription used in the simulations.

4.6.3. Tidal dissipation

The simulations presented in this study considers the dynamics of close-in super-Earths, for which tidal interactions with the central star could lead to significant eccentricity damping. Goldreich and Soter (1966) (see also Jackson et al., 2008) derived the tidal eccentricity damping timescale of the planet, $\tau_{\rm tidal}$ as

$$\tau_{\rm tidal} = \left[\frac{4}{63} \left(G M_\star^3 \right)^{-1/2} \frac{M_{\rm p}}{R_{\rm p}^5} Q_{\rm p} \right] a_{\rm p}^{13/2}, \tag{4.7}$$

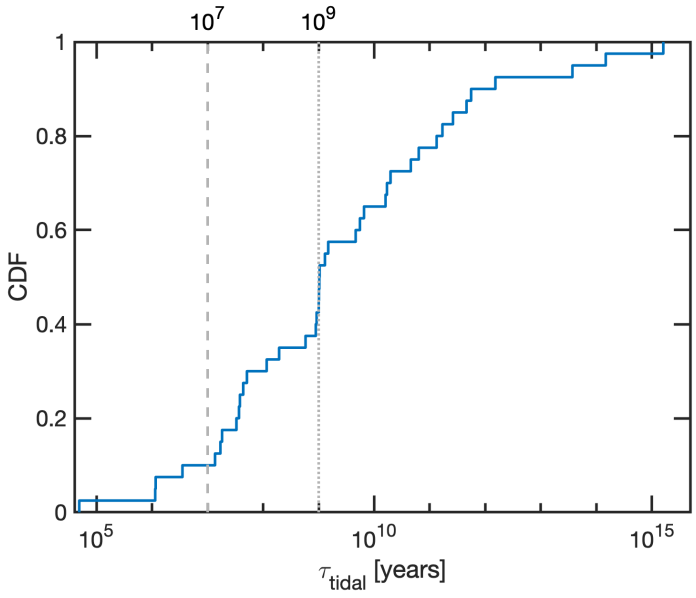

Figure 4.19.: CDFs of the tidal eccentricity damping time-scale obtained by equation 4.7. The vertical dashed line marked the 10 Myr time (main simulation runtime) and the dotted line marked the 1 Gyr time.

where $R_{\rm p}$ is the radius of the planet and $Q_{\rm p}$ is the tidal dissipation parameter. Figure 4.19 shows the values of $\tau_{\rm tidal}$ obtained from equation 4.7 with the value of $Q_{\rm p} = 100$ for all the planets contained in the inner planetary templates. The appropriate value of $Q_{\rm p}$ is uncertain, and is thought to range between $100 \lesssim Q_{\rm p} \lesssim 10^4$, with smaller values applying to Earth-like bodies and the larger values applying to planets with significant gas envelopes. For the adopted value of $Q_{\rm p}$ we see that 10 % of the planets in the inner systems have $\tau_{\rm tidal}$ smaller than the simulation runtime of 10^7 yr, and more importantly $\sim 50\%$ of the planets have $\tau_{\rm tidal} \leq 1$ Gyr, indicating that tides should play an important role over the typical ages of the Kepler systems.

Equation 4.7 shows that the eccentricity damping time is a strong function of semi-major axis, and hence within a multiplanet system we would expect the inner-most planet to experience the strongest tidal damping. Secular interactions combined with tidal dissipation, however, can increase the efficiency with which the eccentricities of more distantly orbiting planets are damped, and hence it is necessary to consider the coupled evolution of entire planetary systems when considering the effects of tidal dissipation on the observed eccentricity distribution. To examine the effect of tides on our results, we have extended the simulations labelled as `6g.60M.5AU`, including tidal forces on all planets operating on timescales given by equation 4.7. Ideally we would run the simulations for 1 Gyr, but this is not possible for the short-period

systems we are considering, so we instead adopt the value of $Q_{\rm p} = 1$ and a runtime of 10 Myr, which should be equivalent to running the simulations for 1 Gyr with $Q_{\rm p} = 100$ because of the linear relationship between $Q_{\rm p}$ and $\tau_{\rm tidal}$.

Figure 4.20 shows the eccentricity distributions for the `6g.60M.5AU` subset of runs, where the upper panel shows the results after the first 10 Myr without tidal forces applied and the lower panel shows the results after another 10 Myr during which tidal damping was applied. The colour for each data point illustrates the fractional change of eccentricity, $\Delta e = (e_{\rm i} - e_{\rm f})/e_{\rm i}$. As expected, tidal damping is more effective for close-in planets ($a \lesssim 0.1$ au). The heavily damped planets (blue data points) contain $\sim 30 - 40$ % of the overall population, which is similar to the CDF value at 10^9 yr from Figure 4.19. These heavily damped planets end up in essentially circular orbits.

Figure 4.21 compares the synthetically observed eccentricity distributions for the systems with and without tidal damping. Around 35 % of the single-transiting planets and 20 % of the planets in multi-transiting systems end up with very low eccentricities ($e < 0.01$), and we see in general that tides cause a significant shift in the observed eccentricity distributions of both single- and multi-transiting systems. A contributing factor in explaining these changes is the fact that the most heavily damped planets are also the innermost planets of the systems, which have the highest probability to transit ($P \sim R_\star/a_{\rm p}$). However, more distantly orbiting planets are also observed, and these are not strongly affected by tides, and so contribute a significant number of high eccentricity planets to the observed distributions. It is interesting to note that a clear eccentricity dichotomy is maintained, even in the presence of eccentricity damping, when comparing the eccentricities of single- and multi-transiting planets.

The shifted eccentricity distributions have mean and median eccentricities of $\{\langle e_1 \rangle, \tilde{e}_1\} = \{0.17, 0.11\}$ and $\{\langle e_{\geq 2} \rangle, \tilde{e}_{\geq 2}\} = \{0.04, 0.04\}$, which are clearly smaller than those obtained from the `6g.60M.5AU` simulations without eccentricity damping: $\{\langle e_1 \rangle, \tilde{e}_1\} = \{0.25, 0.20\}$ and $\{\langle e_{\geq 2} \rangle, \tilde{e}_{\geq 2}\} = \{0.07, 0.06\}$. Hence, these values are significantly influenced by tides, as expected, but also maintain a significant dichotomy between single- and multi-transiting systems. This dichotomy signal is stronger than that which arises from the *in situ* self scattering model (e.g. Poon et al., 2020).

Compared to the Kepler systems examined by Mills et al. (2019), for which $\{\langle e_1 \rangle, \langle e_{\geq 2} \rangle\} = \{0.21, 0.05\}$, we see that the simulations with tides applied now produce moderately smaller values of the mean eccentricities compared to the Kepler systems. We note, however, that Mills et al. (2019) list the seven most eccentric

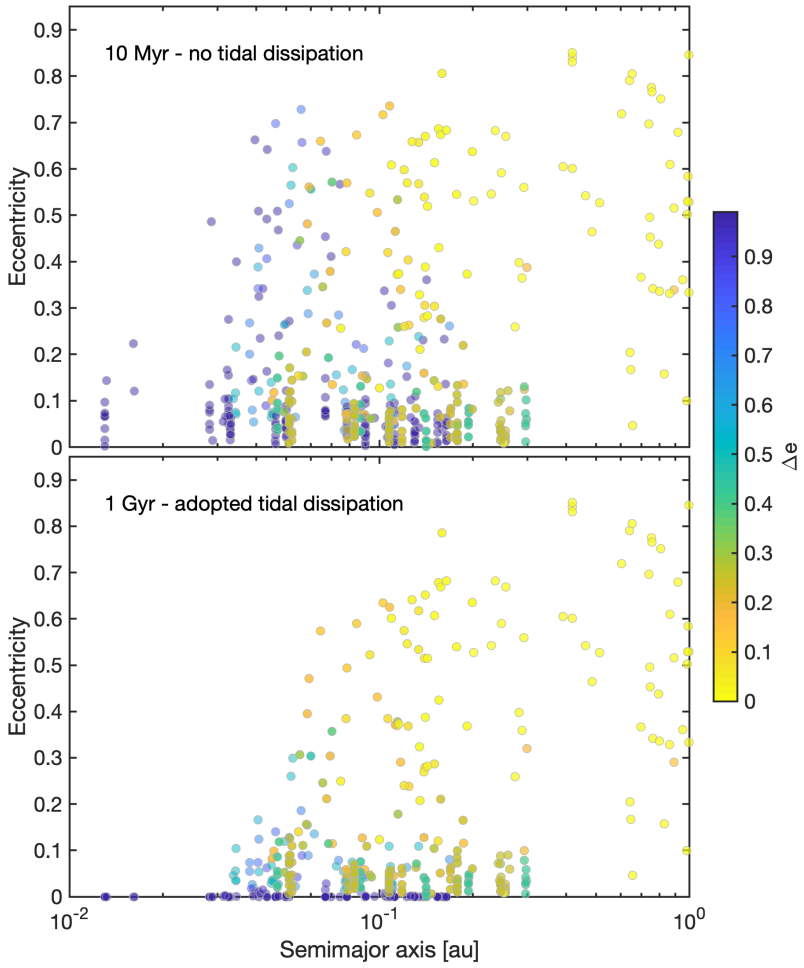

Figure 4.20.: Eccentricities versus semi-major axes of the simulation results from `6g.60M.5AU` template. Upper panel: the simulation results from the main simulations (10 Myr N-body, without tidal dissipation); Lower panel: the simulation results of the next 10 Myr N-body + 1 Gyr adopted tidal evolution. The colour for each data point illustrates the fractional change of eccentricity, Δe.

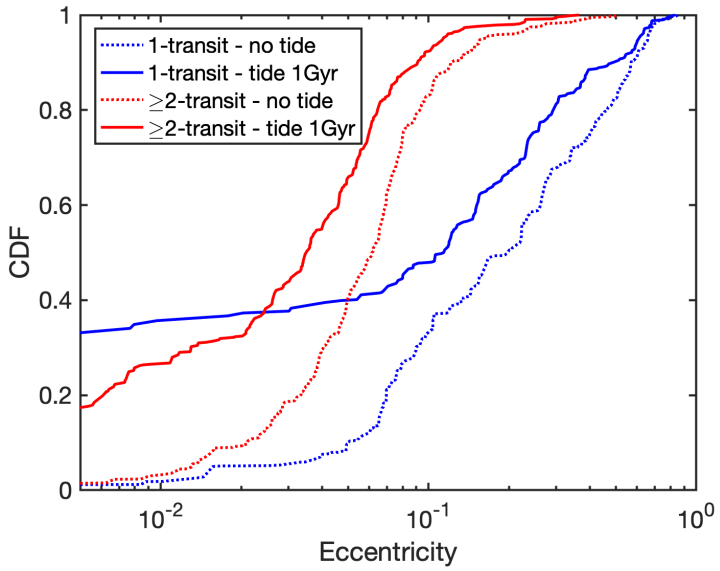

Figure 4.21.: CDFs of the eccentricities obtained from the synthetic observations of the simulations for the `6g.60M.5AU` template by the 10 Myr no tidal damping model (dotted lines) and the next 10 Myr + 1 Gyr tidal damping model (solid lines). The dotted lines are the same as the thick solid lines from figure 4.17 in the same colours.

single-transiting systems in their table 1, and five out the seven have orbital periods in excess of 16 days. Furthermore, six out of the seven planets have radii ≥ 2.2 $R_\oplus$, such that the appropriate value of Q_p may significantly exceed our adopted value of $Q_p = 100$ if these planets have significant gas envelopes. Hence, it seems likely that the most eccentric Kepler systems identified by Mills et al. (2019) have orbital and physical parameters that render tidal eccentricity damping relatively ineffective over Gyr timescales. Mills et al. (2019) consider the possibility that the single-transiting systems are actually composed of two sub-populations: low and high eccentricity systems. Our results support this hypothesis, and show that these two populations can be explained by tidal eccentricity damping.

4.7. Discussion and conclusions

We have presented the results of N-body simulations of outer systems of giant planets that coexist with inner compact systems of super-Earths. The outer planetary systems are set up to become dynamically unstable, and the purpose of this study is to examine whether or not perturbations from the outer planets, acting on the inner systems, are able to generate dichotomies in the multiplicity and eccentricity

distributions that agree with the Kepler data for compact systems of super-Earths.

There have been previous studies of the influence of outer giant planets perturbing inner systems (Mustill et al., 2017; Huang et al., 2017), and in this work we have examined the influence of the multiplicities, orbital radii and masses of the planets that make up the outer systems. We have considered systems of $N_{\rm out} = 3$, 6 and 12 outer planets, that are centred on orbital radii $\tilde{a}_{\rm out} = 2$, 5 and 10 au. The planet masses were varied between $M_{\rm p,out} = 15$ and 100 Earth masses, and we chose parameters that ensure these bodies would be below the detection thresholds of long-term RV surveys. The inner systems were set up using eight of the known 5-planet systems that were discovered by Kepler as templates. We ran 100 simulations for each combination of inner and outer system templates considered.

The final multiplicities of the inner systems are found to be highly dependent on the architectures of the outer systems. The multiplicity tends to be smaller when the initial values of $N_{\rm out}$ and $M_{\rm p,out}$ increase and $\tilde{a}_{\rm out}$ decreases. Some outer system templates fail to induce significant perturbations on the inner systems, leaving the multiplicity unchanged ('unperturbed systems'), while other outer system templates induce very strong perturbations, leading to instabilities in the inner systems in all runs performed for that template.

For outer systems with $N_{\rm out} = 3$ and $\tilde{a}_{\rm out} = 10$ au, the gravitational scatterings among the outer planets are not strong enough for any of the giant planets to come close enough to the inner systems to generate noticeable disturbances. On the other hand, the $N_{\rm out} = 12$ systems generated strong perturbations on the inner systems, and only the subset of runs with $\tilde{a}_{\rm out} = 10$ au allows some systems to survive relatively unperturbed. Except for the `12g.15M.10AU` template, all other $N_{\rm out} = 12$ templates have an ~ 80 per cent chance of completely destroying the inner system, while the other ~ 20 per cent mostly have one inner planet surviving. Hence, the occurrence rates of inner systems of planets that arise from the $N_{\rm out} = 12$ runs are much smaller than implied by the Kepler data for planets around Solar-type stars (Fressin et al., 2013; Petigura et al., 2013). A recent analysis by Zhu et al. (2018) suggests the mean multiplicity of super-Earth systems with periods < 100 days is ~ 3, and approximately 1/3 of Sun-like stars host compact planetary systems. The $N_{\rm out} = 12$ outer systems we have considered produce outcomes that are in clear disagreement with these numbers.

We selected eight outer system templates with $N_{\rm out} = 6$ at $\tilde{a}_{\rm out} = 5$ or 10 au for a more in-depth investigation of their effects on the inner systems. The results show a wide range of final multiplicities, from the template `6g.15M.10AU` resulting in unperturbed inner systems in ~ 75 % of the runs, to `6g.100M.5AU` completely

destroying ~ 75 % of the inner systems. A common outcome is the generation of single planet inner systems with relatively high eccentricities, as required by the observed eccentricity dichotomy. The simulations also produce a clear relation between the final multiplicities and the eccentricity distributions, where inner systems with higher final multiplicities have lower eccentricities. For runs where the final planet number in the inner system $N_{\rm in,final} = 5$, the final eccentricity distribution is similar to the initial distribution, as expected. The $N_{\rm in,final} = 1$ systems, however, have large eccentricities with a mean of $\langle e_1 \rangle = 0.48$.

We undertook synthetic transit observations of the final inner systems. We counted the relative numbers of one-planet, two-planet, ..., five-planet systems detected by the synthetic observations, and compared this to the ratios of the observed multiplicities in the Kepler data. We found that the simulated systems in this chapter produce better agreement with the Kepler data compared to the multiplicity ratios obtained from N-body simulations that only consider self-excitation of inner systems of super-Earths (e.g. chapter 3 and Poon et al., 2020). Some of our outer templates, such as those with $\tilde{a}_{\rm out} = 5$ au, resulted in inner systems that are in very good agreement with the Kepler multiplicity ratios, and reproduce the Kepler multiplicity dichotomy.

Synthetic observation of the simulated inner systems produces a very clear eccentricity dichotomy. The single transiting planet systems always have a significantly higher mean eccentricity, $\langle e \rangle$, than the multi-transit systems. Compared to control runs which were performed for inner systems without any outer giant planets, the differences between the mean eccentricities for single systems, $\langle e_1 \rangle$, and those for multiple systems, $\langle e_{\geq 2} \rangle$, are much larger in our models with outer planets. The control runs produce $\{\langle e_{\geq 2} \rangle, \langle e_1 \rangle\} = \{0.05, 0.08\}$, whereas we obtain $0.06 \leq \langle e_{\geq 2} \rangle \leq 0.1$ and $0.11 \leq \langle e_1 \rangle \leq 0.25$ for the runs with outer systems. The outer system comprised of 6 planets centred around semi-major axis 5 au with masses of 60 $M_\oplus$ resulted in an eccentricity dichotomy characterised by $\{\langle e_{\geq 2} \rangle, \langle e_1 \rangle\} = \{0.07, 0.25\}$, which is in decent agreement with the values $\{\langle e_{\geq 2} \rangle, \langle e_1 \rangle\} = \{0.04, 0.3\}$ reported by Xie et al. (2016) and the values $\{\langle e_{\geq 2} \rangle, \langle e_1 \rangle\} = \{0.05, 0.21\}$ reported by Mills et al. (2019). While this is the outer system template that gives the best agreement with the Kepler data, other templates resulted in similar eccentricity dichotomies. One feature of our study, however, is that the mean eccentricities for multiple transiting systems coming out of the simulations are always larger than those reported for the Kepler data. This may in part be due to our choice of initial conditions for the inner system templates, but it is also influenced by the fact that systems of 2 transiting planets contribute significantly to the overall eccentricity distributions of the multi-

ple transiting systems. The synthetically observed 2 transiting planet systems often come from underlying systems that have been significantly perturbed by the outer systems, such that the mean eccentricity increases above $\langle e_{\geq 2} \rangle = 0.05$.

We ran some additional simulations to test the effect of including additional physics in the models. Simulations that adopted a realistic collision model, instead of the simple hit-and-stick model adopted in the main suite of simulations, produced results that are very similar to the original simulation set. Hence, we conclude that the collision model is not important for determining the outcome of the simulations. We also considered the effect of eccentricity damping due to tidal interaction with the central star over Gyr timescales, and here we observe a significant change in the eccentricity distributions of both the single- and multi-transiting systems. In spite of this change, we still maintain a significant dichotomy between the final eccentricity distributions for the single- and multi-transiting systems, such that the mean eccentricities obtained are $\{\langle e_{\geq 2} \rangle, \langle e_1 \rangle\} = \{0.04, 0.17\}$, similar to but slightly smaller than the values $\{\langle e_{\geq 2} \rangle, \langle e_1 \rangle\} = \{0.05, 0.21\}$ for the Kepler systems obtained by Mills et al. (2019).

Finally, we remark that although we have undertaken a wide ranging parameter study of outer planetary systems influencing inner systems of super-Earths, what we have presented here is far from exhaustive in terms of multiplicity, planet mass and semi-major axes for the outer systems. For example, we have considered outer planetary systems consisting of identical planets in terms of mass and radius, and this might influence the outcome in terms of collisions versus ejections (Anderson et al., 2020). Furthermore, we have also assumed that the inner systems are fully formed at the time when the outer giant planets undergo dynamical instability, and it is possible that in a number of systems the instability occurs earlier during the epoch of formation. These considerations may lead to different outcomes, such that the multiplicity and eccentricity dichotomies can be more accurately reproduced by models than has been achieved in this work. Nonetheless, the study presented here shows that the general scenario of outer unseen planets perturbing inner planetary systems is a promising mechanism for explaining some aspects of the Kepler data.

5. Hot Jupiters with companion super-Earths

從前望永遠 然後看現在
沒再問蝴蝶與莊周卻問時日那可偷
存在共真相 如有兩樣
只有繼續尋訪

— 張敬軒, 林子祥, 潘源良 〈相對論〉

This chapter mainly follows the material that is contained in Poon et al. (submitted). The title of this publication is '*In situ formation of hot Jupiters with companion super-Earths*', and it was submitted for publication in the Monthly Notices of the Royal Astronomical Society.

Observations have confirmed the existence of multiple-planet systems containing a hot Jupiter and smaller planetary companions. Examples include WASP-47, Kepler-730, and TOI-1130. We examine the plausibility of forming such systems *in situ* using N-body simulations that include a realistic treatment of collisions, an evolving protoplanetary disc and eccentricity/inclination damping of planetary embryos. Initial conditions are constructed using two different models for the core of the giant planet: a 'seed-model' and an 'equal-mass-model'. The former has a more massive protoplanet placed among multiple small embryos in a compact configuration. The latter consists only of equal-mass embryos.

Simulations of the seed-model lead to the formation of systems containing a hot Jupiter and super-Earths. The evolution consistently follows four distinct phases: early giant impacts; runaway gas accretion onto the seed protoplanet; disc damping-dominated evolution of the embryos orbiting exterior to the giant; a late chaotic phase after dispersal of the gas disc. Approximately 1% of the equal-mass simulations form a giant and follow the same four-phase evolution. Synthetic transit observations of the equal-mass simulations provide an occurrence rate of 0.26% for systems containing a hot Jupiter and an inner super-Earth, similar to the 0.2%

occurrence rate from actual transit surveys, but simulated hot Jupiters are rarely detected as single transiting planets, in disagreement with observations. A subset of our simulations form two close-in giants, similar to the WASP-148 system. The scenario explored here provides a viable pathway for forming systems with unusual architectures, but not for the majority of hot Jupiters.

5.1. Introduction

Since the discovery of the hot Jupiter 51 Pegasi b (Mayor and Queloz, 1995), observations using the radial velocity (RV) technique have unveiled a large number of such objects (e.g. Udry et al., 2002; da Silva et al., 2006; Butler et al., 2006; Quinn et al., 2012; Flagg et al., 2019).[1] The population study by Mayor et al. (2011) indicates that $\sim 14\%$ of Sun-like stars host a gas giant planet with mass $\gtrsim 50$ $M_\oplus$ and orbital period $\lesssim 5000$ days, and the fraction of stars hosting such planets with periods $\lesssim 10$ days is $\sim 1\%$.

Various scenarios have been proposed to explain the origins of hot Jupiters, including formation at larger distances followed by migration driven by the gaseous protoplanetary disc, *in situ* formation after the build-up of a large reservoir of planetary building blocks close to the star, and migration due to tidal interaction with the central star after a giant planet achieves a high eccentricity orbit due to planet-planet scattering, or through the Kozai-Lidov effect induced by a distant companion (see Dawson and Johnson (2018) and references therein for a comprehensive discussion on the origins of hot Jupiters).

Transit surveys have also discovered hot Jupiters. Of particular interest are the discoveries of multiple systems that contain a transiting hot Jupiter and inner planetary companions, such as WASP-47 (Becker et al., 2015), Kepler-730 (Cañas et al., 2019), and TOI-1130 (Huang et al., 2020). These multiple planet systems cannot have formed through high-eccentricity migration, and hence must have formed *in situ* or after large scale migration. Previous studies have examined the consequences for planet formation of a giant planet migrating over large distances, and have shown that a natural outcome can be the formation of relatively compact multi-systems consisting of hot Jupiters and nearby super-Earths/sub-Neptunes (e.g. Fogg and Nelson, 2005, 2007; Raymond et al., 2006; Mandell et al., 2007). Our aim in this study is examine whether or not such systems naturally arise from an *in situ* formation scenario.

[1] All exoplanet data used in this chapter are from the NASA Exoplanet Archive unless stated otherwise.

One argument deployed in favour of the migration scenario for the formation of hot Jupiters is that a protoplanetary disc with mass characteristic of the Minimum Mass Solar Nebula model (MMSN, Hayashi, 1981) lacks sufficient mass in solids and gas to build a planet *in situ*. Instead, it is envisaged that planet cores form beyond the snow line and start to accrete their gas envelope at large distances from the star (Pollack et al., 1996). Subsequent angular momentum exchange with the disc leads to inwards migration (e.g. Lin et al., 1996; Tanaka et al., 2002; Nelson and Papaloizou, 2004; Alibert et al., 2005; Kley and Nelson, 2012; Coleman and Nelson, 2014, 2016a; Bitsch et al., 2019), during which accretion of gas onto the planet continues. Various arguments have been used against this scenario. For example, Bailey and Batygin (2018) suggest that the inner envelope of the mass-period distribution for hot Jupiters is best explained by a model in which hot Jupiters accrete their gas *in situ*. This argument hinges on both the planet mass and the size of the magnetospheric cavity in a protoplanetary disc depending on the mass accretion rate through the disc. It is possible a migration based scenario might also produce such a relation if a migrating planet accretes gas through a gap at a rate that scales with the overall accretion rate through the disc. Differences in the planetary mass distributions when comparing hot and cold giants, where the mean mass of hot Jupiters is slightly lower than that of cold Jupiters, also seems to be in tension with the migration scenario, since this trend is not expected if hot and cold Jupiters both formed at large distance.

Transit surveys such as Kepler have discovered a large population of close-in planets (Borucki et al., 2010, 2011; Batalha et al., 2013; Burke et al., 2014; Rowe et al., 2015; Mullally et al., 2015; Coughlin et al., 2016; Thompson et al., 2018), and have shown that systems of multiple super-Earths/sub-Neptunes with orbital period < 100 days are common around Sun-like stars (e.g. Fressin et al., 2013; Dressing and Charbonneau, 2013, 2015; Petigura et al., 2013; Zhu et al., 2018). Some of these compact planetary systems contain a significant amount of mass (e.g. Lissauer et al., 2011a), and a number of studies have been conducted to examine *in situ* formation scenarios, where it is assumed a large amount of solid mass has drifted into the inner disc before accumulating into planetary building blocks which then collide to form compact planetary systems (e.g. Chiang and Laughlin, 2013; Hansen and Murray, 2013; Moriarty and Ballard, 2016; Matsumoto and Kokubo, 2017; Mustill et al., 2018; Poon et al., 2020; Matsumoto et al., 2020, and chapter 3). It is then a natural extension of this scenario to suppose that hot Jupiters can form *in situ* if a sufficiently massive core forms during the gas disc life time such that it can accrete a massive envelope (e.g. Lee et al., 2014; Boley et al., 2016; Batygin et al., 2016;

Hasegawa et al., 2019).

In this chapter we use N-body simulations to study the *in situ* formation of planetary systems containing a hot Jupiter and at least one inner planetary companion. We assume there is sufficient solid material in the inner region of the protoplanetary disc (a total of ~ 25 to 30 $M_\oplus$ at $\lesssim 0.5$ au) to build a core that can undergo runaway gas accretion. The solid mass is distributed among multiple ($N > 50$) protoplanets which dynamically evolve and collide to build the final planetary system, and if a planet grows to be large enough during the gas disc life time then it is able to accrete gas and become a giant planet. We adopt two sets of initial conditions, namely the 'seed-model' and the 'equal-mass-model'. The seed-model initially contains a more massive seed-protoplanet that ensures a hot giant can form in the simulation. This biased initial condition allows us to study the dynamical history and behaviour of the gas giant *in situ* model. The equal-mass-model initially contains equal mass protoplanets, where numerous protoplanet mergers in the early stages of a simulation are essential to form a giant planet. This model allows us to investigate the occurrence rate of giant planets in the *in situ* model in an unbiased manner.

The possibility of significant gravitational scattering of protoplanets during the growth of a giant planet, combined with the large Keplerian velocities close to the star, have motivated us to adopt a realistic collision algorithm in our N-body simulations (Leinhardt and Stewart, 2012). We employ a 1D evolving viscous disc model, which is used to implement eccentricity and inclination damping forces onto the protoplanets. Migration torques, however, are neglected in this study. Our model includes gas accretion onto planetary cores when they become sufficiently massive, gap opening, and photoevaporation of the disc. The realistic collision model allows us to track the nature of planet-planet collisions, and the generation of collisional debris during high energy impacts, and we use this information to compare the collisional behaviour between the gas-rich and gas-free phases. The giant formation rate is sensitive to the gas envelope accretion model adopted, and we examine the effect of this on the outcomes of our simulations.

This chapter is structured as follows. In Section 5.2, we describe the simulation methods, including the realistic collision model and the disc model. In Section 5.3, we describe the initial conditions. In Section 5.4, we present the results and analyses of our simulations. In Section 5.5, we investigate the impact of assuming different gas envelope accretion models. Finally, we discuss our results and draw conclusions in Section 5.6.

5.2. Simulation model

We use the N-body code SYMBA (Duncan et al., 1998) to undertake the simulations presented in this chapter, and we utilise a version that implements an algorithm to treat planet-planet collisions realistically. We also include an evolving protoplanetary disc model to provide a prescription for planet-disc interactions, which we describe below.

5.2.1. Realistic collision model

The Leinhardt and Stewart (2012) collision model was implemented in SYMBA as described our previous study (Poon et al., 2020, see also section 2.1). Here we just summarise the post-collision outcomes that can arise in the simulations. For a more detailed description we refer the reader to the above cited papers (see also Kokubo and Genda, 2010; Genda et al., 2012; Stewart and Leinhardt, 2012; Scora et al., 2020).

We refer to the more massive body involved in a collision as the target, and the less massive object as the projectile. Post collision, we can have a largest remnant, a second largest remnant and debris in the form of superplanetesimals. Superplanetesimals are non-mutually interacting gravitating particles with masses between 0.1 to 10 times the mass of Ceres. The radii of all objects are calculated using the mass-radius relation equation A.6. The collision algorithm consists of a decision tree with the following possible outcomes that depend on how the impact velocity, $V_{\rm imp}$, compares with the escape velocity, $V_{\rm esc}$:

1. Perfect merger, when $V_{\rm imp} < V_{\rm esc}$. A single body is formed with the total mass and momentum of the original two collided bodies.

2. Supercatastrophic disruption, when $V_{\rm imp}$ exceeds the threshold for supercatastrophic disruption. No massive bodies remain and all the mass is in the form of collision debris represented by superplanetesimals.

3. Catastrophic disruption, when $V_{\rm imp}$ exceeds the threshold for catastrophic disruption. Only one massive body remains and the rest of the mass is in the form of collision debris.

4. Erosion, when $V_{\rm imp}$ exceeds the threshold for erosion. The target is eroded by the projectile and loses some of its mass, while the projectile is completely disrupted into superplanetesimals.

5. Partial accretion, when $V_{\rm imp}$ is less than the threshold for erosion and the collision angle is smaller than the critical angle. The target gains mass from the projectile, which is completely disrupted into numerous superplanetesimals.

6. Hit-and-spray, when the collision angle is larger than the critical collision angle and $V_{\rm imp}$ exceeds the velocity of transition to hit-and-spray. The projectile is completely disrupted into debris.

7. Hit-and-run, similar to the hit-and-spray criterion but with $V_{\rm imp}$ between the velocities for transition to hit-and-spray and hit-and-run. The projectile mass is reduced and the remaining mass goes into debris.

8. Bouncing collision, similar to the hit-and-run criterion but with $V_{\rm imp}$ smaller than the velocity for transition to hit-and-run and greater than the normalised critical impact velocity (Kokubo and Genda, 2010; Genda et al., 2012). The target and projectile retain all of their mass and the collision is treated as an inelastic bounce.

9. Graze-and-merge, similar to bouncing criterion but with $V_{\rm imp}$ smaller than the normalised critical impact velocity. The outcome is similar to perfect merge where a single body forms containing all the mass of the colliding objects.

When a supercatastrophic collision happens, it can lead to the formation of a stable ring composed of superplanetesimals that sits interior to the system of planetary embryos. This can be very stable and remain present throughout the simulations. To avoid integrating the large number of debris particles using a small time-step size, we adopted the mass reduction scheme described in Poon et al. (2020) and appendix A.1, which is based on determining the collision frequency of the debris particles and assuming collisions will grind the debris to dust, which is then quickly removed from the system.

5.2.2. Protoplanetary disc model

Here we describe the viscously evolving protoplanetary disc model and the eccentricity/inclination damping forces that are applied to the planetary embryos (see also section 2.2).

Disc evolution

The disc adopted here is a typical α-disc model (Shakura and Sunyaev, 1973). The kinematic viscosity, ν, is given by

$$\nu = \alpha c_{\rm s}^2 \Omega^{-1} = \alpha \left(\frac{k_{\rm b} T}{\mu m_{\rm H}} \right) \left(\frac{G M_\star}{r^3} \right)^{-\frac{1}{2}}, \tag{5.1}$$

where $c_{\rm s}$ is the local sound speed, Ω is the local angular velocity, T is the local temperature of the disc, μ is the mean molecular weight of the disc and set as 2.4 u in this study, $m_{\rm H}$ is the atomic mass of hydrogen, G is the gravitational constant, $M_\star$ is the mass of the host star and r is the local distance from the host star. We set $\alpha = 10^{-3}$.

The evolution of the disc surface density, Σ, is given by the usual diffusion equation, augmented by additional terms (equation 2.48) that account for the torque exerted by a gap forming planet (if present in a simulation, see also equation 2.41) and photoevaporation:

$$\frac{\partial \Sigma}{\partial t} = \frac{1}{r} \frac{\partial}{\partial r} \left[3 r^{1/2} \frac{\partial}{\partial r} \left(\nu \Sigma r^{1/2} \right) - \frac{2 \Lambda \Sigma r^{3/2}}{\left(G M_\star \right)^{1/2}} \right] - \frac{\partial \Sigma_{\rm pe}}{\partial t}, \tag{5.2}$$

where Λ is the injection rate of angular momentum per unit mass due to tidal interaction with a planet (Lin and Papaloizou, 1986), and $\partial \Sigma_{\rm pe} / \partial t$ is the rate of change of the disc surface density due to photoevaporation (see equation 5.5). Λ is given by

$$\Lambda = {\rm sign}(r - r_{\rm p}) \frac{G M_\star}{2r} \left(\frac{M_{\rm p}}{M_\star} \right)^2 \left(\frac{r}{\max(H, |r - r_{\rm p}|)} \right)^4, \tag{5.3}$$

where $r_{\rm p}$ is the distance between the planet and the host star, $M_{\rm p}$ is the mass of the planet, and H is the local disc scale height obtained from $H = c_{\rm s} / \Omega$. Equation 5.3 is only applied when a planet mass reaches the gap opening criterion, otherwise $\Lambda = 0$. Gap opening occurs when the following criterion is satisfied (Crida et al., 2006):

$$50 \alpha \left(\frac{M_\star}{M_{\rm p}} \right) \left(\frac{H}{r_{\rm p}} \right)^2 + \frac{3}{4} \left(\frac{3 M_\star}{M_{\rm p}} \right)^{1/3} \left(\frac{H}{r_{\rm p}} \right) - 1 \leqslant 0. \tag{5.4}$$

During the late stage of a disc's lifetime, when the accretion rate is low, photoevaporation due to UV radiation from the star starts to dominate the disc surface density evolution (e.g. Clarke et al., 2001; Matsuyama et al., 2003; Ruden, 2004; Takeuchi et al., 2005). The photoevaporation term, $\partial \Sigma_{\rm pe} / \partial t$ in equation 5.2, is

obtained using the prescription in Dullemond et al. (2007):

$$\frac{\partial \Sigma_{\rm pe}}{\partial t} = 1.16 \times 10^{-11} \Phi_{41}^{1/2} r_{\rm g}^{-3/2} S_{\rm g} \left(\frac{\rm M_\odot}{\rm au^2 yr} \right), \tag{5.5}$$

where $r_{\rm g}$ is the gravitational radius, $S_{\rm g}$ is a scaling factor in terms of $r_{\rm g}$, and Φ_{41} is the rate of emitted ionizing photons from the host star in units of 10^{41} photons $\rm s^{-1}$. The scaling factor $S_{\rm g}$ is calculated according to

$$S_{\rm g} = \begin{cases} \left(\dfrac{r}{r_{\rm g}} \right)^{-2} \exp \left(\dfrac{r - r_{\rm g}}{2r} \right) & \text{for } r \leq r_{\rm g}, \\ \left(\dfrac{r}{r_{\rm g}} \right)^{-5/2} & \text{otherwise.} \end{cases} \tag{5.6}$$

Hollenbach et al. (1994) simply defined $r_{\rm g}$ to be where the sound speed equals the Keplerian orbital velocity. In this study, we set $r_{\rm g} = 10$ au and $\Phi_{41} = 1$.

Eccentricity and inclination damping

A planet orbiting in a gaseous disc will experience eccentricity and inclination damping forces. It may also experience torques that drive migration, but in this study we neglect these and examine how systems evolve only under the influence of eccentricity/inclination damping, while remaining agnostic about why migration does not occur.

We follow Papaloizou and Larwood (2000) in our implementation of the damping forces. The accelerations of an object due to the eccentricity damping, $d\boldsymbol{v}_{\rm e}/dt$, and inclination damping, $d\boldsymbol{v}_{\rm i}/dt$, are given by

$$\frac{d\boldsymbol{v}_{\rm e}}{dt} = -2 \frac{v_{\rm r}}{t_{\rm e}} \hat{\boldsymbol{r}}, \tag{5.7}$$

and

$$\frac{d\boldsymbol{v}_{\rm i}}{dt} = -2 \frac{v_{\rm z}}{t_{\rm i}} \hat{\boldsymbol{z}}, \tag{5.8}$$

where $v_{\rm r}$ is the radial velocity, $v_{\rm z}$ is the velocity in the $\boldsymbol{z}$ direction (perpendicular to the disc plane), $t_{\rm e}$ is the eccentricity damping timescale, and $t_{\rm i}$ is the inclination damping timescale.

We use the expressions from Cresswell and Nelson (2008) to calculate $t_{\rm e}$ and $t_{\rm i}$, giving

$$t_{\rm e} = \frac{t_{\rm wave}}{0.780} \left[1 - 0.14 \left(\frac{e}{h} \right)^2 + 0.06 \left(\frac{e}{h} \right)^3 + 0.18 \left(\frac{e}{h} \right) \left(\frac{I}{h} \right)^2 \right], \tag{5.9}$$

and

$$t_{\rm i} = \frac{t_{\rm wave}}{0.544}\left[1 - 0.30\left(\frac{I}{h}\right)^2 + 0.24\left(\frac{I}{h}\right)^3 + 0.14\left(\frac{I}{h}\right)\left(\frac{e}{h}\right)^2\right], \tag{5.10}$$

where e is the eccentricity, I is the inclination, h is the ratio between the disc scale height and the local radius (H/r), and $t_{\rm wave}$ is the characteristic time of the orbital evolution. Tanaka and Ward (2004) defined $t_{\rm wave}$ as

$$t_{\rm wave} = \left(\frac{M_{\rm p}}{M_\star}\right)^{-1}\left(\frac{\Sigma_{\rm p} a_{\rm p}^2}{M_\star}\right)^{-1}\left(\frac{c_{\rm s}}{a_{\rm p}\Omega_{\rm p}}\right)^4 \Omega_{\rm p}^{-1}, \tag{5.11}$$

where $a_{\rm p}$ is the semi-major axis of the object, and the values of $\Sigma_{\rm p}$, $c_{\rm s}$, and $\Omega_{\rm p}$ are evaluated at $a_{\rm p}$.

The damping timescales $t_{\rm e}$ and $t_{\rm i}$ obtained from equations 5.9 and 5.10, are applied until a planet satisfies the gap opening criterion (equation 5.4), after which we set the damping timescales to be 20 local orbital periods.

5.2.3. Gas envelope accretion

We allow gas to accrete onto protoplanets if their masses reach a threshold value. We adopt different gas accretion prescriptions to examine the sensitivity of our results to the assumed gas accretion rates, and these prescriptions are described below.

Simple model

In our simple model, a protoplanet can start to accrete a gaseous envelope from the protoplanetary disc when its mass reaches 3 $\rm M_\oplus$. We adopt the expression for the gas accretion rate from Coleman and Nelson (2016b), which provides a fit to the 1D calculations presented in Movshovitz et al. (2010) for a planet located at 5.2 au from its host star:

$$\frac{dM_{\rm ge}}{dt} = \dot{M}_0\left(\frac{M_{\rm core}}{\rm M_\oplus}\right)^{2.4} \exp\left[\frac{M_{\rm ge}}{\rm M_\oplus}\left(\frac{1}{22} + \frac{\rm M_\oplus}{M_{\rm core}}\right)\right]. \tag{5.12}$$

Here $\dot{M}_0$ is the mass accretion rate for a 1 $\rm M_\oplus$ protoplanet core (with no gas envelope), $M_{\rm core}$ is the mass of the protoplanet core, $M_{\rm ge}$ is the mass of the accreted gas envelope, and we have the relation $M_{\rm p} = M_{\rm core} + M_{\rm ge}$. The value of $\dot{M}_0$ used is that in Coleman and Nelson (2016b), namely $\dot{M}_0 = 4.656 \times 10^{-8}$ $\rm M_\oplus yr^{-1}$.

Gas accretion based on fits to local 1D models

We also run a subset of simulations using an alternative local gas accretion model (see section 5.5). For this model, we calculated new fits to gas accretion rates obtained using the 1D envelope structure model of Coleman et al. (2017) (see also Papaloizou and Terquem, 1999; Papaloizou and Nelson, 2005). Whilst the gas accretion rates used in Coleman and Nelson (2016b) were based on fits from a handful of simulations in Movshovitz et al. (2010), they do not take into account the local disc properties, i.e. surface density and temperature. This was appropriate for those models since the simulations performed in Movshovitz et al. (2010) were based on formation scenarios for Jupiter located at 5.2 au.

In taking the local disc properties into account, we ran a large number of simulations where we placed planetary cores of masses between $2-15$ $\mathrm{M}_{\oplus}$ at orbital radii spanning $0.2-50$ au in the evolving disc model of Coleman and Nelson (2016b). Using the gas envelope and accretion model of Coleman et al. (2017), these cores were able to accrete gas until the protoplanetary disc dispersed, or until they reached a critical state where runaway gas accretion occurs and they become giant planets. We included an opacity reduction factor f_{opa} that reduces the grain opacity contribution, similar to other works (e.g. Mordasini et al., 2014).

As the discs were evolving, the local disc properties for the planets were ever-changing and this was taken into account when calculating the 1D gas envelope structure. These local disc properties were recorded, as well as the planet properties (e.g. core mass, envelope mass) and the mass accretion rate. In total the simulations provided $\sim 50,000$ data points to determine a gas accretion rate, a significant improvement on the 20 data points used to formulate the equations in Coleman and Nelson (2016b). With these results we were able to fit an equation that takes into account not only the properties of the protoplanets, but also the local properties of the disc, notably the temperature, T_{local}, and the opacity reduction factor, f_{opa}. The local gas accretion rate can then be calculated according to

$$\begin{aligned}\left(\frac{dM_{\mathrm{ge}}}{dt}\right)_{\mathrm{local}} =& 10^{-10.199}\left(\frac{\mathrm{M}_{\oplus}}{\mathrm{yr}}\right) f_{\mathrm{opa}}^{-0.963}\left(\frac{T_{\mathrm{local}}}{1\ \mathrm{K}}\right)^{-0.7049} \\ &\times \left(\frac{M_{\mathrm{core}}}{\mathrm{M}_{\oplus}}\right)^{5.6549}\left(\frac{M_{\mathrm{ge}}}{\mathrm{M}_{\oplus}}\right)^{-1.159} \\ &\times \left[\exp\left(\frac{M_{\mathrm{ge}}}{M_{\mathrm{core}}}\right)\right]^{3.6334}. \end{aligned} \tag{5.13}$$

The strong dependence on the core mass in equation 5.13 is notable, as is the depen-

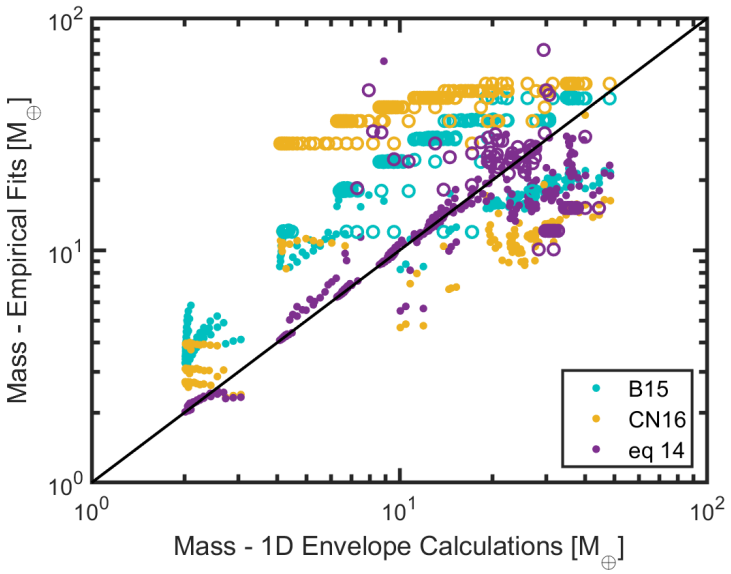

Figure 5.1.: Comparison of final planet masses from gas accretion fits to those calculated using the 1D gas envelope structure model of Coleman et al. (2017). The different colour markers represent different accretion fits arising from Bitsch et al. (2015, B15 equation 17), Coleman and Nelson (2016b, CN16), and Equation. 5.13. The open circle markers indicate where the planets have reached the runaway gas accretion phase, and would become giant planets. The black line indicates a 1:1 relation between the final planet masses obtained using the fits and the 1D model.

dence of the gas envelope mass fraction. Such strong dependences are in agreement with Coleman et al. (2017) where the accretion rate is heavily dependent on the core and envelope masses. The dependence on the temperature is also consistent with previous works where gas accretion rates decrease as the local disc temperature rises. Note that the effect of surface density is not included in this equations, since it is found to have a very weak effect.

When comparing the planet masses attained using equation 5.13 to those arising from the 1D envelope structure model, we find that they are in good agreement. Figure 5.1 shows the final planet masses arising from the 1D envelope structure model (x-axis) versus those attained from different gas accretion fits (y-axis). The purple points correspond to using equation 5.13, the yellow points come from equation 5.12 (Coleman and Nelson, 2016b, CN16), and the cyan points arise from the fits in Bitsch et al. (2015, B15) to the work of Piso and Youdin (2014).

To calculate these masses, we ran numerous single-planet-in-disc models where we placed cores of different masses at different locations in the disc, and allowed

them to accrete locally. We did not include any migration in these models, and we stopped each model when the planets reached either: a critical state where the planet envelope can no longer be hydrostatically supported (1D models), a gas accretion rate of 2 $M_\oplus$/1000 yr (CN16 model, equation 5.13), or where the envelope mass was equal to double the core mass (the B15 model). We denote planets that reach a critical state using open circles, whereas closed circles denote planets that did not reach the relevant criterion. In all of these scenarios, a period of runaway gas accretion is expected to occur with the planets becoming giant planets.

The final planet masses obtained using equation 5.13 (purple markers) sit close to the black diagonal line, that represents a 1:1 ratio between the planet masses found through either the 1D envelope structure model or the gas accretion fits. This is especially true for low mass planets, i.e. super-Earth mass range. The only region where the fits become less consistent with the 1D model is at higher planet masses, where the fits reach a critical state slightly earlier or later than occurs in the 1D model. This will result in only small differences in the final planet masses, since these occurrences happen early in the disc lifetime, and the planets in all cases will undergo runaway gas accretion and become giant planets.

When comparing the results from other works to the 1D structure models, for the higher mass planets it is clear that like the masses found through equation 5.13, the planets reached a critical state slightly before/after the 1D model, where again these planets would undergo runaway gas accretion and become giant planets. Where the results from other works disagree with the 1D model is at lower planet masses particularly between 4 $M_\oplus \leq M_\mathrm{p} \leq$ 15 $M_\oplus$. For both accretion fits (CN16 and B15), the planets typically reached a critical state, whereas the planets in the 1D models did not. This would ultimately lead to the C16 and B15 fits producing an over abundance of giant planets compared to the 1D model results.

For much lower mass planets, $M_\mathrm{p} \leq 4$ $M_\oplus$, the accretion rates from Bitsch et al. (2015) and Coleman and Nelson (2016b) do not reach a critical state. This is more consistent with the 1D models, where the planets do not have large enough core masses to be able to accrete significant gaseous envelopes. However, even though there is greater agreement between the 1D models and the fits here, the masses arising from the accretion rate fits are still considerably larger than those arising from the 1D models. This is in contrast to the masses arising from equation 5.13 where there is excellent agreement with the 1D models at low planet mass.

Accretion after gap opening

The local gas accretion model (equation 5.12 or 5.13) applies until a planet satisfies the gap opening criterion (equation 5.4) or when a protoplanet is sitting in a gap which was opened by another massive protoplanet. The accretion rate then changes to

$$\left(\frac{dM_{\rm ge}}{dt}\right)_{\rm gap} = \min\left[\frac{3\pi\nu_{\rm vsd}\Sigma_{\rm vsd}}{N_{\rm p,gap}}, \left(\frac{dM_{\rm ge}}{dt}\right)_{\rm local}\right], \tag{5.14}$$

where $N_{\rm p,gap}$ is the number of protoplanets that sit in the gap and are accreting gas through viscous supply, $\nu_{\rm vsd}$ is the local viscosity and $\Sigma_{\rm vsd}$ is surface density in the disc exterior to the gap. To ensure that the numerator $3\pi\nu_{\rm vsd}\Sigma_{\rm vsd}$ measures the typical accretion rate through the disc, and not the local value in the gap, we evaluate it at 2 au from the star.

In figure 5.2 we demonstrate the difference between applying equation 5.12 and 5.13 when calculating the gas accretion rate, including the switch to equation 5.14 once a gap has been opened. The final masses obtained when adopting equation 5.13 shows a higher sensitivity to the local disc parameters compared to the model applying equation 5.12. This mainly arises due to the increase in disc midplane temperature closer to the star, reducing the gas accretion onto the planets and having a much larger effect for smaller core masses.

For the main simulation results and analysis presented in section 5.4, we use the accretion rates provided by equations 5.12, while in section 5.5 we apply equation 5.13 for comparison.

5.3. Simulation set-up

In this study, we are interested in systems of hot Jupiters with coexisting inner super-Earths. We have selected a number of systems that have a hot Jupiter and an inner companion to provide templates for the initial conditions of the simulations. The consideration of system selection is straight forward and uses the following criteria: (i) the system contains a confirmed transiting giant with orbital period less than 30 days; (ii) the system contains at least one transiting companion with orbital period shorter than the orbital period of the giant. There are five systems that meet these criteria. In order of increasing orbital period of the giant they are WASP-47,

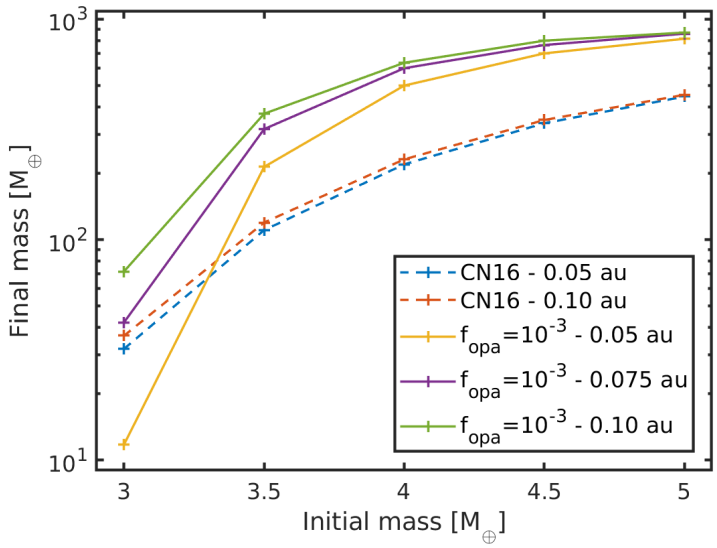

Figure 5.2.: Comparison between the initial and final masses (at the point of gas dissipation) of the planet in the single-planet simulations, the difference is based on the adoption of different gas envelope accretion models and initial orbital radii. The results of applying the Coleman and Nelson (2016b, CN16) routine (equation 5.12) are plotted using dashed lines, where the protoplanet is initially located at 0.05 au (blue) or 0.10 au (orange). The solid lines represent the results obtained when adopting our locally-dependent envelope accretion routine (equation 5.13) with the value of $f_{\rm opa} = 10^{-3}$, where the protoplanet is initially located at 0.05 au (yellow), 0.075 au (purple), or 0.10 au (green). It is clear that the CN16 routine is less dependent on the initial protoplanet mass and orbital location (local disc conditions).

Kepler-730, TOI-1130, Kepler-487[2,3], and Kepler-89 (see figure 5.3, and also table A.4 for basic stellar and planetary parameters of the selected systems).

5.3.1. Template construction

We constructed two sets of templates for our simulations: the 'seed-model' and the 'equal-mass model'. The seed-model contains one higher mass protoplanet, that acts as a seed for the growth of a giant planet, and multiple equal-mass smaller bodies. The equal-mass-model only contains planetary embryos of equal mass. The

[2]The two inner companions, KOI-191.02 and KOI-191.03, of Kepler-487 are candidates instead of confirmed planets.

[3]This system is included because both objects have a relatively low probability to be due to any of the considered astrophysical false positive scenarios (1.4×10^{-3} for KOI-191.02 and 4.2×10^{-5} for KOI-191.03).

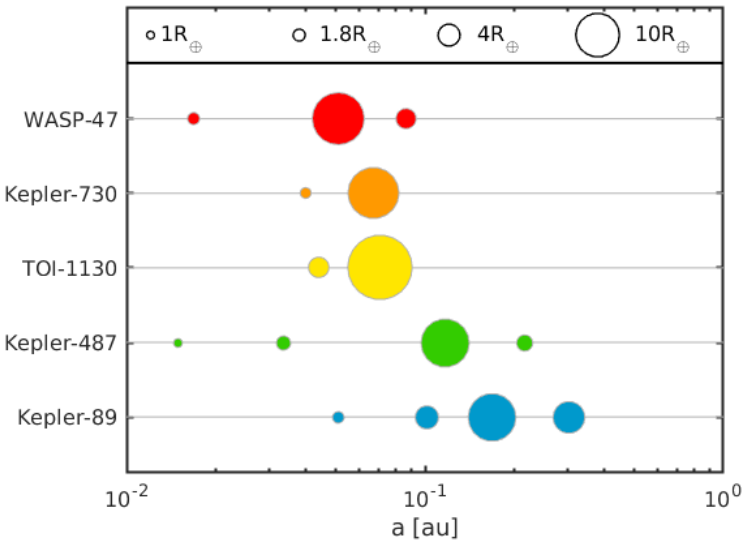

Figure 5.3.: The five selected planetary systems that meet the selection criteria (section 5.3). The symbol size represents the radius of each planet.

configurations for the two templates are:

1. seed-model: 25 inner equal-mass bodies + 1 seed protoplanet + 25 outer equal-mass bodies,

2. equal-mass-model: 51 equal-mass bodies.

All the equal-mass bodies have mass 0.5 $M_\oplus$. There are two subsets of the seed-model, where the mass of the seed protoplanet is either 4.0 or 4.5 $M_\oplus$. The middle body for each template (the 26th body) is located at the semi-major axis of the hot Jupiter from the selected system. The mutual separations between each adjacent body are $K = 5$ for the seed-model, and two subsets of $K = 4$ and 5 for the equal-mass model. This K-value is measured in units of the mutual Hill radius, R_H, which is defined by

$$K = \frac{a_{i+1} - a_i}{R_{\mathrm{H},i}}, \tag{5.15}$$

where

$$R_{\mathrm{H},i} = \frac{a_i + a_{i+1}}{2} \left(\frac{M_{\mathrm{p},i} + M_{\mathrm{p},i+1}}{3M_\star} \right)^{\frac{1}{3}}, \tag{5.16}$$

and subscript i denotes the value of the i-th body in the system.

We use a labelling convention based on the parameters mentioned above when describing the simulations as follows: '(selected system)-(subset)'. In total, there are 20 different simulation templates from the combinations of 5 selected systems (`WASP47`,

`Kepler730`, `TOI1130`, `Kepler487`, and `Kepler89`) and 4 different subsets (`4M-5K`, `4.5M-5K`, `XM-4K`, and `XM-5K`). For example, the `WASP47-4M-5K` template refers to the seed-model runs with a 4 $M_\oplus$ seed-protoplanet (which is located at 0.0513 au) and mutual separation $K = 5$ for all pairs of bodies, and `Kepler730-XM-4K` refers to the equal-mass model runs where the 26-th body is located at 0.0694 au with a mutual separation $K = 4$ for all pairs of embryos. The central stars of each system have their masses and radii taken from table A.4.

Each template is run with 10 different instances of the initial conditions. The initial eccentricities of the bodies are randomly drawn from a Rayleigh distribution with eccentricity parameter, $\sigma_e = 2 \times 10^{-3}$. The initial inclinations for each run are randomly drawn from a Rayleigh distribution with inclination parameter, $\sigma_I = 1 \times 10^{-3}$ rad. The distributions follow the relation $e = 2I$ (where I is in radians), but the initial values of e and I for each planet are independent. The arguments of pericentre, ω, longitudes of ascending node, Ω, and mean anomalies, M are distributed uniformly in the range $0 \leq (\omega, \Omega, M) < 2\pi$. Objects whose orbital distances exceed 100 au are removed from the simulations. The time steps used are set to be 1/20th of the shortest orbital period in the system. Each simulation is run for 10^7 yr.

5.3.2. Disc parameters

The protoplanetary disc model (section 2.2 and 5.2.2) is included in all N-body simulations. The disc surface density profile, $\Sigma_{\rm init}(r)$, has the same power-law index as the MMSN (Hayashi, 1981):

$$\Sigma_{\rm init}(r) = \Sigma_{\rm init}(1\,{\rm au})\, r^{-1.5}, \tag{5.17}$$

where $\Sigma_{\rm init}(1\,{\rm au}) = 1700$ g cm^{-2} is the initial surface density of the disc at 1 au. The temperature profile is fixed throughout the simulations and also follows the power index of the MMSN and is given by

$$T(r) = T(1\,{\rm au})\, r^{-0.5}, \tag{5.18}$$

where $T(1\,{\rm au}) = 280$ K is the temperature at 1 au. For the main simulations (section 5.4), the gas envelope accretion routine uses equation 5.12. For the investigation of the impact from the gas envelope accretion routine (section 5.5), the gas envelope accretion rate is given by equation 5.13.

The inner and outer boundaries of the disc are located at 0.05 and 30.0 au, re-

spectively. We apply a zero radial velocity condition at the outer boundary, which ensures that no additional mass flows into the disc model. We apply a zero torque condition at the inner boundary to allow accretion onto the star. Any bodies sitting outside of the boundaries of the disc do not interact with it, and so experience no eccentricity/inclination damping and do not accrete gas.

5.4. Results

5.4.1. Evolution of the seed-model

We now present the results of simulations that explore the *in situ* formation of systems containing a hot Jupiter and inner super-Earths. To recap, two sets of simulations were performed for the seed-model, one with the seed protoplanet mass equal to 4.0 $M_\oplus$ (`4M-5K`), and the other with 4.5 $M_\oplus$ (`4.5M-5K`).

Dynamical evolution

The dynamical evolutions for the seed-models can be classified in to four distinct phases in terms of their behaviour. For convenience when discussing the different phases, we will refer to them as the *early impact*, *runaway gas accretion*, *outer disc damping*, and *late chaotic phase*.

The early impact phase corresponds to the time early in the simulations when the embryos undergo dynamical instability and collisional accretion, before any planet in the system is massive enough to trigger its runaway gas envelope accretion (figure 5.4, the first ~ 2 Myr). The collisions between protoplanets in this stage are induced by the small initial mutual separations, and generally take place early in the simulations (< 1 Myr for initial $K = 5$). After the initial impacts have occurred, the disc damping forces start to dominate the orbital evolution, and the eccentricities and inclinations are forced to remain very low. Except for the seed body, it is difficult to form a planet with mass greater than 4 $M_\oplus$ because the damping forces prevent the occurrence of mergers.

As the name suggests, the runaway gas accretion phase starts when a planet enters runaway gas accretion (figure 5.4, orange shaded area). The planet increases its mass exponentially in this stage and dynamically heats up the system. The dynamical heating by the swift increase in mass outstrips the disc damping forces, and orbit crossing of the bodies is a common outcome. This phase normally lasts for only a few hundred thousand years. Once the body enters the runaway accretion phase,

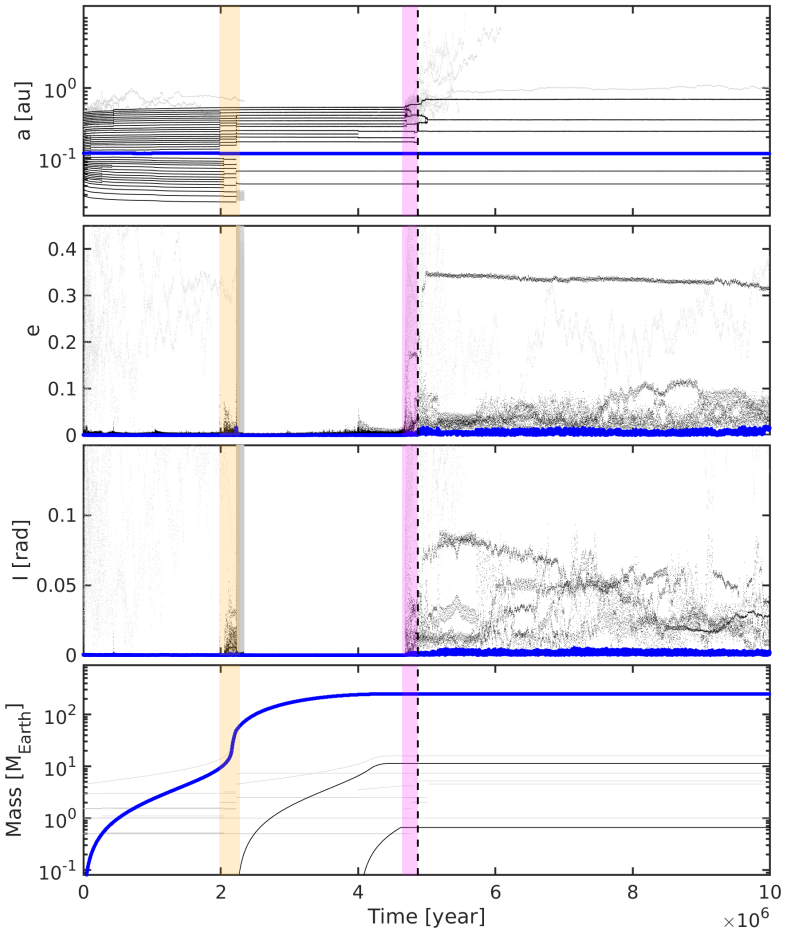

Figure 5.4.: Dynamical history of a run from the `Kepler487-4.5M-5K` template, including the semimajor axes (top panel), eccentricities (second panel), inclinations (third panel), and planet masses (bottom panel). The seed-protoplanet is marked in blue, other protoplanets are marked in black, and collision debris are marked in grey. The black dashed vertical line denotes the point of disc dissipation. The orange shaded area marks the beginning of the runaway gas accretion phase. The magenta shaded area denotes the transition between the end of outer disc damping phase and the beginning of the late chaotic phase.

especially in the inner parts of the discs that we consider in this study, it can open a gap and accrete gas at the viscous supply rate (equation 5.14). This slows down the growth rate of the seed protoplanet and ends the runaway gas accretion phase.

The outer disc damping phase ensues after gap opening in the gas disc, and is illustrated in figure 5.4 by the temporal domain lying between the orange and magenta shaded regions. In the parameter space that we consider in the seed-model, the planet-induced gap sits at an orbital radius < 0.15 au. The inner part of the disc (in terms of the gap location) has a much shorter collisional accretion time compared to the outer part of the disc (disc beyond the gap). This phase is dynamically cooler than the previous runaway phase. The bodies in the inner part of the disc experience almost no disc damping forces, but the impacts experienced during the runaway phase result in relatively large mutual separations developing between the inner protoplanet pairs, so the inner system can be relatively stable. Inner protoplanets can grow to masses comparable to the initial seed mass, and start to accrete their own gas envelopes. However, these large cores initiate gas envelope accretion too late to become gas giants before the gas disc dissipates. The disc damping forces are still a dominating influence in the outer part of the disc. Similar to the early impact stage, protoplanets mergers are not very common during this phase, and generally we do not find that multiple massive cores form during the disc life time in the region sitting exterior to the giant planet's orbital radius.

The evolution enters the late chaotic phase when photoevaporation dominates the gas surface density evolution. In this stage, the density of the disc is low, and the damping forces gradually become negligible, until the disc eventually dissipates (figure 5.4, magenta area and onward). Self-scattering between protoplanets can heat up the system dynamically, causing dynamical instabilities and giant impacts. Protoplanets can only accrete through giant impacts in this gas-free environment. This stage lasts until a long-term stable system emerges.

Throughout the evolution, the seed protoplanet retains a low eccentricity and inclination (figure 5.4, blue line). The seed also essentially preserves its initial semi-major axis, and generally experiences just a $\sim 1\%$ decrease. For example, in figure 5.4, the initial semimajor axis is 0.117 au and its final value is 0.116 au. This is due to exchanges of energy and angular momentum with the surrounding bodies, combined with the disc damping forces.

We noticed that our seed for the `WASP47-4M` and `WASP47-4.5M` templates did not produce a hot giant in any of the runs for the above reason. The initial semimajor axes (0.051 au) of these seeds are located very close to the inner edge of the disc (0.050 au), and they stop accreting gas once the seeds move out of the disc. To

form a hot giant orbiting with a semimajor axis like WASP-47b might require a seed initially at ~ 0.055 au for our model set-up.

Figure 5.5 shows some additional instances from the suite of runs shown in figure 5.4 (i.e. from the `Kepler487-4.5M-5K` template), together with the evolution of the disc. The upper panel shows the initial conditions, where the blue line is the initial surface density (equation 5.17) and the green dots are the protoplanets. The second panel shows a moment at the beginning of the runaway gas accretion phase, where the seed has accreted enough mass to open a gap in the disc. The third panel shows the moment during the outer disc damping phase. By this time the inner super-Earths/sub-Neptunes were already formed while the outer protoplanets are still in a compact configuration due to the disc damping. The bottom panel shows the moment after the gas disc dissipates. Due to the lack of the disc damping forces, giant impacts between the outer embryos are common and allow the sub-Earth mass protoplanets to grow to super-Earths. Figure 5.6 shows a similar result to figure 5.5, but for a run from the `Kepler730-4M-5K` template. The `Kepler730` templates follow similar behaviour to the `Kepler487` templates, and provide similar final architectures of the planetary systems. The giants formed in the `Kepler730` runs are closer-in at $\sim$ 0.05 to 0.06 au, similar to the WASP-47, Kepler-730, and TOI-1130 systems, where all the giants formed have orbital periods less than 10 days.

Multiplicities, masses, and orbital parameters

There is a systematic difference, in terms of multiplicities and mass distributions, between the final inner and outer parts of the systems from the seed-model (see table 5.1). The inner multiplicity, $N_{\rm in}$, always has a lower value than the outer multiplicity, $N_{\rm out}$. The inner total mass, $M_{\rm total,in}$, is similar to the outer total mass, $M_{\rm total,out}$, in the `Kepler730` and `TOI1130` templates, while the `Kepler487` and `Kepler89` templates hold a relation of $M_{\rm total,in} > M_{\rm total,out}$. This difference is because of the initial conditions of the `Kepler730` and `TOI1130` templates located most of the inner protoplanets between the host star and the inner edge of the disc, while the `Kepler487` and `Kepler89` templates are further out and allow most of the protoplanets to sit inside the disc when $K = 5$. The difference between the two sets of $M_{\rm total,in}$, to a certain extent, shows how much gas mass the inner systems can accrete from the disc. A rough relation of $M_{\rm total,in} \approx 2M_{\rm total,out}$ can be drawn, which denoted a $\sim 1:1$ solid-to-gas ratio for the inner systems. The inner systems also show a higher average mass of the planets, $\langle M_{\rm total,in} \rangle \,/\, \langle N_{\rm in} \rangle$, compared to the average planet mass of

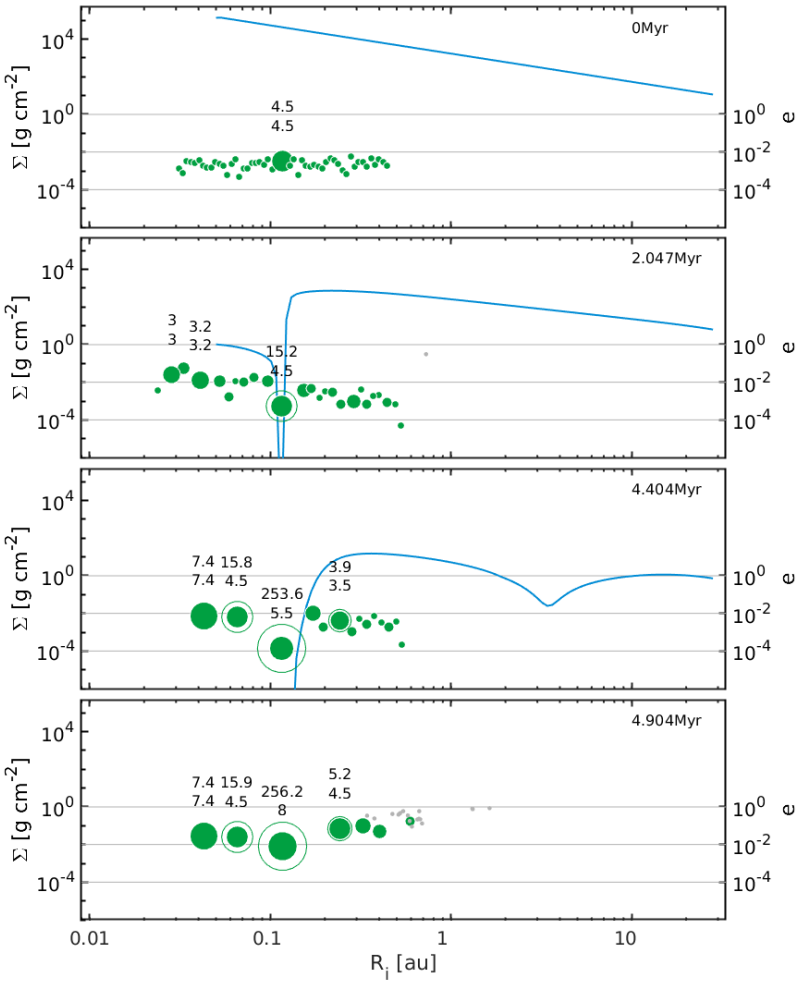

Figure 5.5.: An example formation history from an instance of the the seed-model template `Kepler487-4.5M-5K` (same template as shown in figure 5.4). The blue lines are the disc surface density, green dots are the protoplanets, green circles are the gas envelope and grey dots are the collision debris. Sizes of the greens dots and circles are denoted by their relative mass. Protoplanet masses that are greater than 3 $M_\oplus$ are shown by the text (upper: total mass; lower: core mass). Eccentricities are shown in the right vertical axes. The top panel shows the initial conditions, the second panel is the time that the first gap opens in the disc ($\sim$ 2.0 Myr), the third panel is the time when photoevaporation start to dominate the surface density evolution ($\sim$ 4.4 Myr), and the bottom panel is the time after the disc has dissipated ($\sim$ 4.9 Myr).

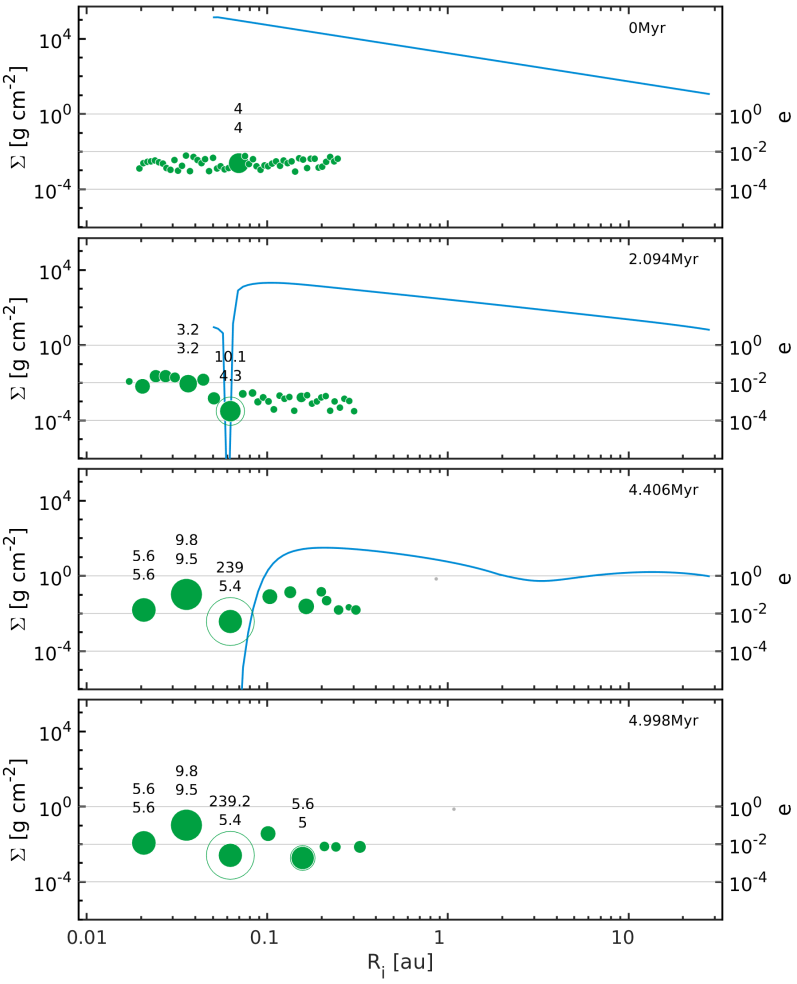

Figure 5.6.: Similar to figure 5.5 but for another seed-model run from the template of `Kepler730-4M-5K`. A closer-in giant planet is formed at $\sim$ 0.05 to 0.06 au, similar to the initial semimajor axis of the seed-protoplanet.

the outer systems, $\langle M_{\mathrm{total,out}} \rangle$ / $\langle N_{\mathrm{out}} \rangle$. The inner average masses of the `Kepler487` and `Kepler89` templates are comparable to Neptune's mass, and the template with closer-in seeds, `Kepler730` and `TOI1130`, are more likely to host super-Earths.

There are no systematic differences in the final orbital parameters when comparing the `4M-5K` and `4.5M-5K` runs, but a clear divergence emerges between the giants and other planets. As shown in figure 5.4, 5.5 and 5.6, the giants retain low eccentricities throughout the simulations. The final mean eccentricity of the giants has a value of $\langle e \rangle \approx 0.01$, while the other planets yield $\langle e \rangle \approx 0.06$. And the final mean inclination for the giants is $\langle I \rangle \approx 0.005$ rad and $\langle I \rangle \approx 0.03$ rad for the companions. The final outcomes of the simulations show a strong correlation between the distribution of the eccentricities and inclinations (figure 5.7). This shows that the systems have

Table 5.1.: Final planetary multiplicities and masses for the seed-model runs. The subscripts 'in' and 'out' represent the inner part and the outer part of the system with respect to the location of the giant (the seed). $\langle M_{\rm g} \rangle$ is the average mass of the final giants, where the `4M` runs always give lower values than the `4.5M` runs, comparable to the final mass obtained in the single-planet case (figure 5.2). `WASP47` runs are not included here because no giant was formed at the end of the simulations, for the reason described in section 5.4.1. The table is separated according to the orbital radius of the giant planet, with the 4 upper templates having giants with $a < 0.1$ au and the 4 low templates having $a > 0.1$ au.

Template	$\langle N_{\rm in} \rangle$	$\langle N_{\rm out} \rangle$	$\langle M_{\rm total,in} \rangle$	$\langle M_{\rm total,out} \rangle$	$\frac{\langle M_{\rm total,in} \rangle}{\langle N_{\rm in} \rangle}$	$\frac{\langle M_{\rm total,out} \rangle}{\langle N_{\rm out} \rangle}$	$\langle M_{\rm g} \rangle$
`Kepler730-4M`	1.9	3.9	13.4	20.6	7.1	5.3	297.9
`Kepler730-4.5M`	1.7	3.2	16.8	16.2	9.9	5.1	360.2
`TOI1130-4M`	2.4	3.4	15.8	13.7	6.6	4.0	221.7
`TOI1130-4.5M`	2.4	4.3	13.8	15.7	5.8	3.7	360.9
`Kepler487-4M`	2.4	2.9	34.0	16.5	14.2	5.7	178.6
`Kepler487-4.5M`	2.0	3.6	44.6	17.0	22.3	4.7	258.3
`Kepler89-4M`	2.2	2.8	45.3	14.9	20.6	5.3	174.0
`Kepler89-4.5M`	2.5	3.9	47.4	17.7	19.0	4.5	247.5
Overall	2.2	3.5	28.9	16.5	13.2	4.8	262.4

undergone dynamical relaxation (Kokubo, 2005), as expected.

5.4.2. Comparison to the equal-mass models

Two sets of simulations were performed for the equal-mass model, one with mutual separations $K = 4$ (`XM-4K`), and one with $K = 5$ (`XM-5K`). Without a seed-protoplanet in the simulations, the formation of a giant planet is expected to occur less frequently than in the seed-model. Forming a gas giant requires a sufficiently massive planet to form early in the gas disc life time so it has time to undergo runaway gas accretion before the disc disperses, and this clearly requires a planet to undergo numerous mergers.

Figure 5.8 shows an example of a run that forms a giant planet. A large core with 4.1 $M_\oplus$ forms and enters the runaway gas accretion stage at ~ 2.4 Myr, so this protoplanet has enough time to undergo runaway growth and transition to accreting gas at the viscous supply rate for ~ 2 Myr, eventually becoming a gas giant. On the other hand, figure 5.9 demonstrates a case where no gas giant forms. The largest protoplanet in this run only enters the runaway gas accretion phase and opens a gap at ~ 4 Myr, which is too close to the end of the disc life time for it to form a Jovain

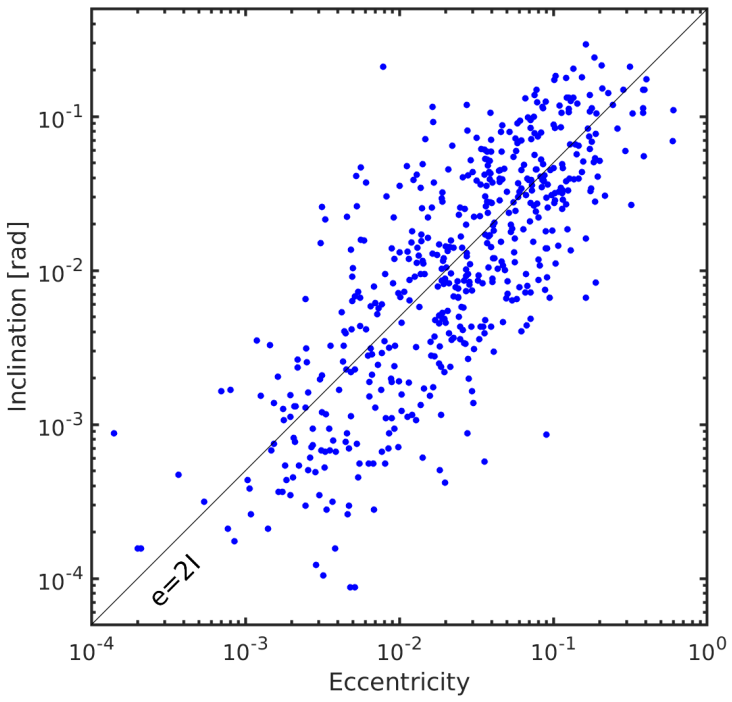

Figure 5.7.: The relation between the final eccentricities and inclinations from the seed-model. The black line represent the relation of $e = 2I$. The eccentricity and inclination distributions provide a good fit to the $e = 2I$ relation, showing a signal of dynamical relaxation.

mass planet.

The formation of gas giants is not common in our equal-mass models. Only 1% of our equal-mass runs (all of them from the $K = 4$ runs) produce a gas giant with final mass greater than 100 $M_\oplus$. We note that the giant formation percentage is sensitive with the gas accretion rate (equation 5.12), and we will discuss the effect of considering different gas accretion prescriptions in section 5.5.

The dynamical evolution observed in the equal-mass models are, in general, different from the evolution history of the seed-model (section 5.4.1), except for the runs in which a giant was formed. For those few runs with giant formation, they follow the dynamical evolution path of the seed-model. More commonly, the equal-mass runs evolve in three phases: *early giant impacts*, *disc damping*, and the *late chaotic phase*. The first and final phases are the same as the seed-model. The disc damping phase is the extension of the early impact phase, where the damping of the eccentricities and inclinations is the dominant effect on the orbital evolution of the protoplanets.

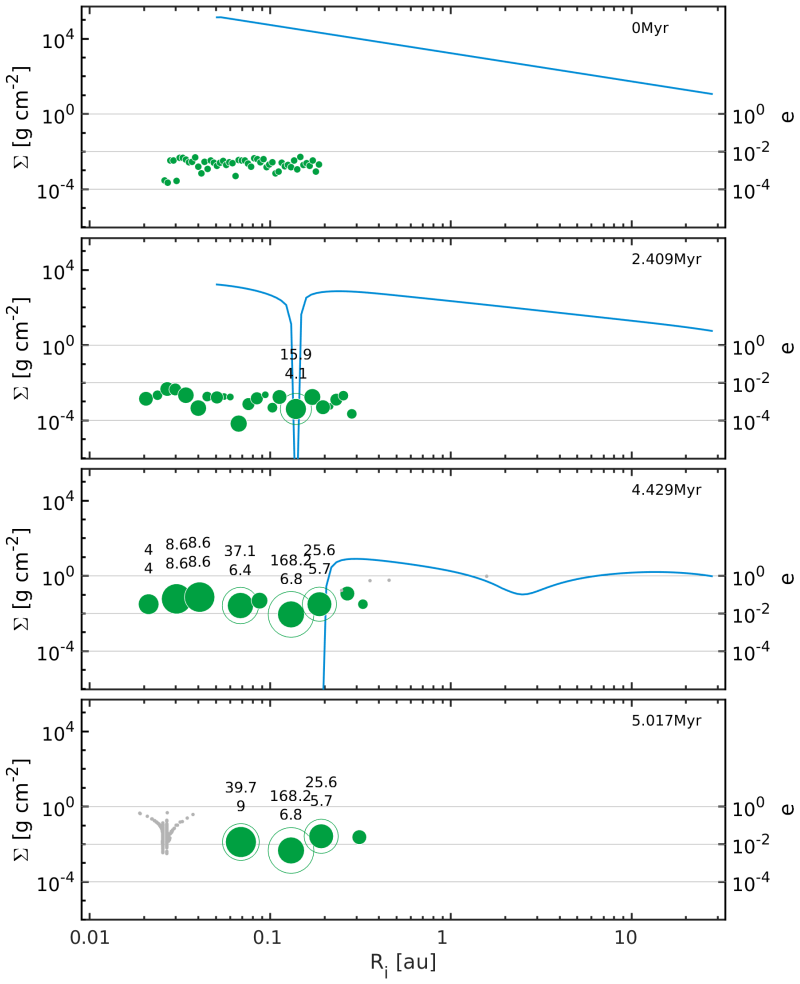

Figure 5.8.: Similar to figure 5.5 but for a run of the equal-mass `Kepler730-XM-4k` template. The gap opening, photoevaporation, and disc dissipation times are similar to the seed-model runs. Unlike the seed-model run (figure 5.6), the formation of giant occurs at a different location to the reference location of the `Kepler730` giant planet.

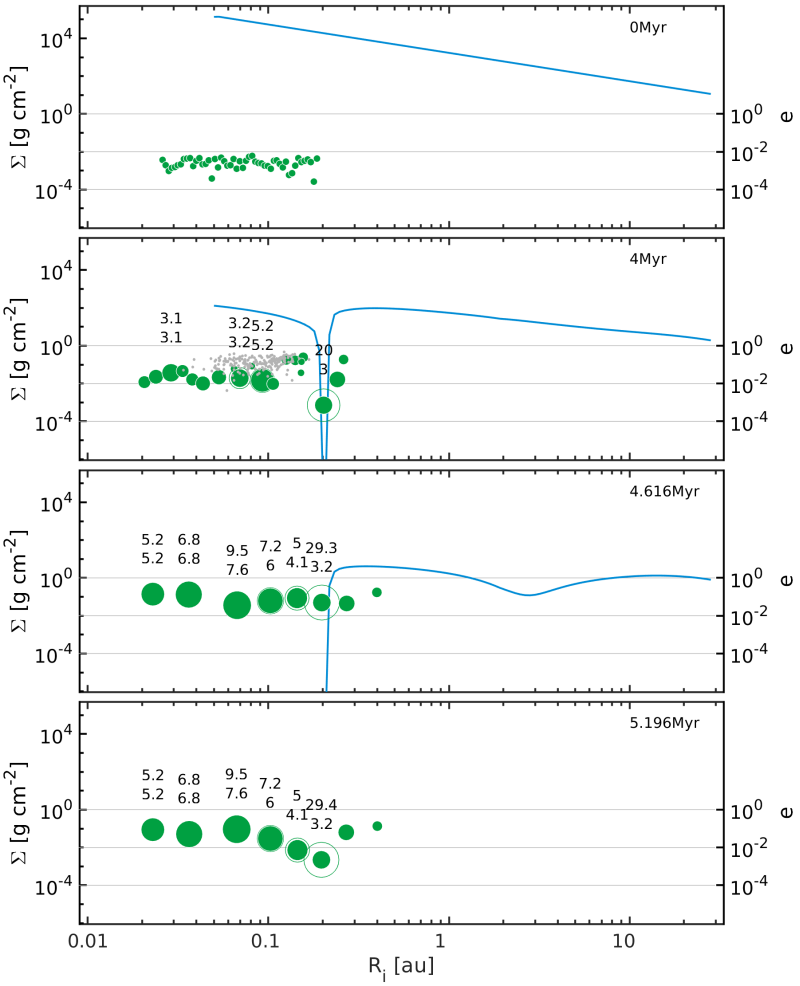

Figure 5.9.: Similar to figure 5.5 and 5.8 but for another run of the `Kepler730-XM-4K` template. This equal-mass run does not result in the formation of a giant, which is a common outcome in the equal-mass models. The transition to runaway gas accretion and gap opening occurs at ~ 4 Myr, leaving insufficient time for the protoplanet to accrete a massive envelope.

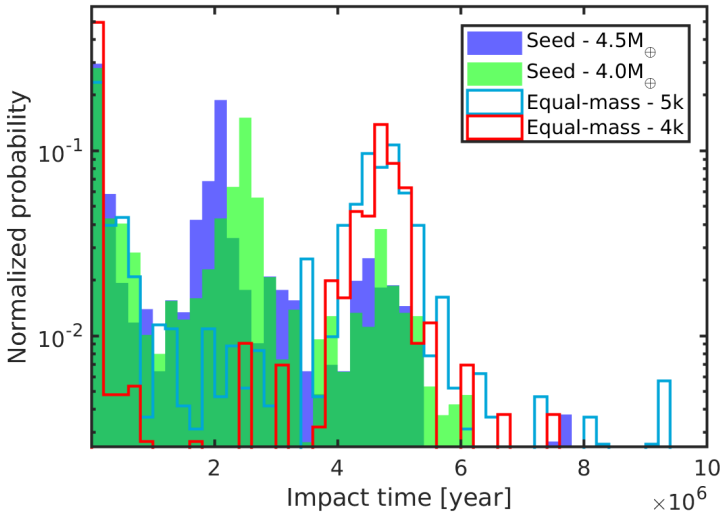

Figure 5.10.: Normalized probabilities of all the giant impact events with respect to time during the four different set of simulations, including the `4.5M` seed-model (blue area), `4M` seed-model (green area), `5k` equal-mass-model (cyan line), and `4k` equal-mass-model (red line).

Figure 5.10 shows the comparison of the impact time between the four subsets of our seed-model and equal-mass model. It is clear that all four of our subsets have an early impact phase. The $K = 5$ systems follow a similar rate of decrease at the number of impacts (< 1 Myr). The $K = 4$ systems (red line) show the majority of early impacts happened before 0.2 Myr, as an effect of a more compact configuration initially. In the `XM-4K` runs, the equal-mass protoplanets can merge with each other to grow sufficiently massive to accrete a noticeable gas envelope earlier than the `XM-5K` runs. This explains why the `XM-4K` runs, compared to the `XM-5K` runs, are more likely to form a giant planet.

The seed-model runs (`4M-5K` and `4.5M-5K`) have peaks between ~ 2.0 to 2.4 Myr, corresponding to the runaway gas accretion phase. This peak comes earlier for the `4.5M-5K` set (blue filled bars) at 2.0 Myr than the `4M-5K` set (green filled bars) at 2.4 Myr. This is simply because the higher mass seeds undergo runaway gas accretion earlier than the lower mass seeds.

All four sets have a peak in the collision times at ~ 5 Myr, which is the average lifetime of our disc models. This marks the onset of the late chaotic phase, and arises because the eccentricity and inclination damping forces diminish as the disc disperses. The peaks for the equal-mass sets are higher than the seed sets. This results from the extended disc damping phase for the equal-mass models, where

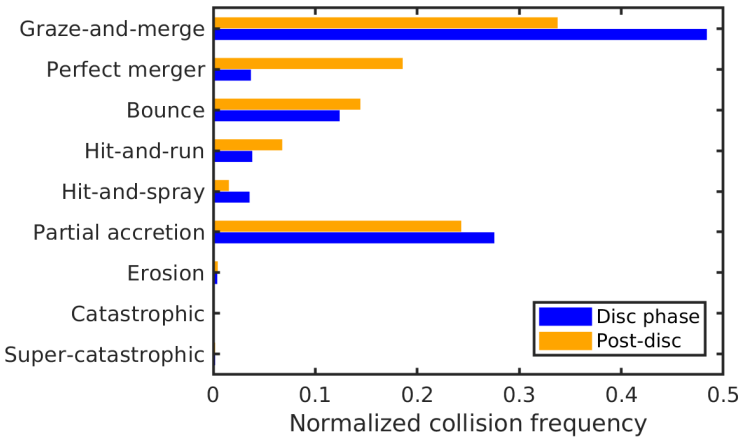

Figure 5.11.: The frequency of each collision type during the two different phases of the simulations. The collisions happening during the disc phase are recorded in the blue bars. The collisions occurring after disc dispersal are recorded in orange bars.

giant impacts are much more common after the disc disperses instead of during the lifetime of the disc. During the gas-free stage, the collisions are more frequent within a million years after the disc has dissipated, and the systems became more dynamically quiet at ~ 6 Myr and beyond.

5.4.3. Collision behaviour

The outcomes of protoplanet collisions show some differences when they occur during the gas disc phase and after the disc dispersed. Figure 5.11 reveals the frequency of different types of collisions. The blue bars indicate the collision frequency during the disc phase, and the orange bars shows the frequency after the disc is no longer present.

In general, the type of collision that can create a large amount of debris, such as super-catastrophic, catastrophic, and erosion collisions, are not common. For the post-disc phase, the most common type of collision is graze-and-merge, followed by partial accretion. The collision frequency results for the post-disc phase is in strong agreement with the gas-free simulations by Poon et al. (2020) and Scora et al. (2020), where these two studies also considered the same collision model (see also chapter 3).

The collision behaviour during the disc phase is different, and almost half of the collisions are graze-and-merge. Together with the drop in the number of perfect

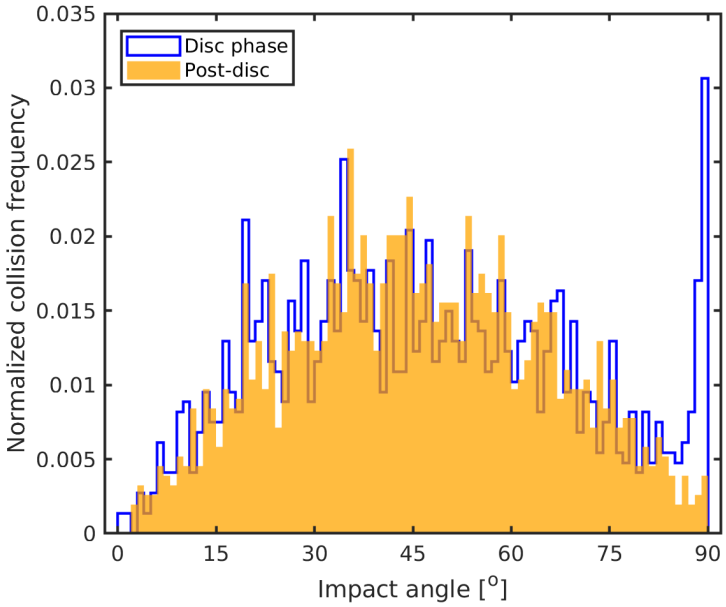

Figure 5.12.: The frequency of impact angles of all the giant impact events in our simulations. The collisions happening during the disc phase are plotted in blue. The collisions occurring after disc dispersal are plotted in orange.

mergers, it indicates that there are more slow collisions with high impact angles compared to the gas-free stage. This phenomenon can also be noticed in figure 5.12, where the figure illustrates the frequencies of different impact angles. The distribution of impact angles for the post-disc stage (orange filled bars) follow a distribution peaking at around 45° (similar to figure 3.16, see also Poon et al., 2020; Scora et al., 2020). Meanwhile, the impact angles in the disc phase (blue line) follow a similar distribution, except in the large angle domain. There is a sharp increase of the frequency at impact angles between 85° to 90°. This indicates that the disc influences the dynamics and induces slow mergers between protoplants at high impact angle. The high impact angles show that the orbits of the two colliding bodies are more circular, instead of arising from high eccentric orbital crossing.

5.4.4. Observational detection rate of hot Jupiter and inner super-Earth systems

Transit surveys have discovered $\sim$ 2500 planetary systems. Among this population, five systems contain a transiting giant with orbital period less than 30 days and transiting super-Earths/sub-Neptunes (section 5.3) that orbit interior to the giant.

These systems are WASP-47, Kepler-730, TOI-1130, Kepler-487, and Kepler-89. These systems make up $\sim 0.2\%$ of the whole population of transiting systems. To compare our simulation outcomes with this detection rate, we carried out synthetic transit observations of our final planetary systems.

Each simulated planetary system is synthetically observed from 100,000 randomly chosen viewing locations, isotropically distributed with respect to each host star. To compare our results to transit surveys, we only consider planets that satisfy the observation limits of a Kepler-like survey. Therefore, we only consider planets, and exclude all collision debris, with orbital radius less than 1 au.

All our seed-model runs (except `WASP47`) contain a hot Jupiter at the end of the simulations. Synthetic transit observations of the final seed-model system tell us that 30.2% of planetary systems that are detected contain a hot Jupiter and an inner super-Earth. In this model, when a giant planet is detected, there is only a $\sim 5\%$ chance that the inner companion will not also be picked up as a transiting planet.

The equal-mass runs also show there is only a $\sim 5\%$ chance that an inner companion will not be detected when a giant is detected in synthetic observations (figure 5.13), similar to the seed-model. Hence, figure 5.13 shows that a prediction of these *in situ* formation simulations is that hot Jupiters detected in transit surveys should almost always be detected with interior super-Earths, and they should be detected with exterior about 50% of the time, which is clearly not the case. Hence the initial conditions we have adopted, or some other aspects of the model, do not apply to the majority of hot Jupiter systems.

To recap, approximately 1% of the equal-mass simulations produce a giant planet (section 5.4.2). Synthetic transit observations of all of our equal-mass simulations show that the proportion of detected planetary systems containing both a hot Jupiter and an inner super-Earth is 0.26%. This is similar to the occurrence rate of $\sim 0.2\%$ from the actual transit observations. There are five systems (as we defined and selected) containing a hot Jupiter and an inner transiting planet within a total number of ~ 2500 transiting exoplanet systems. Nonetheless, in spite of this apparent agreement it is clear the model presented here is not in agreement with the majority of hot Jupiter systems observed in transit surveys that are also able to detect super-Earths, as shown in figure 5.13.

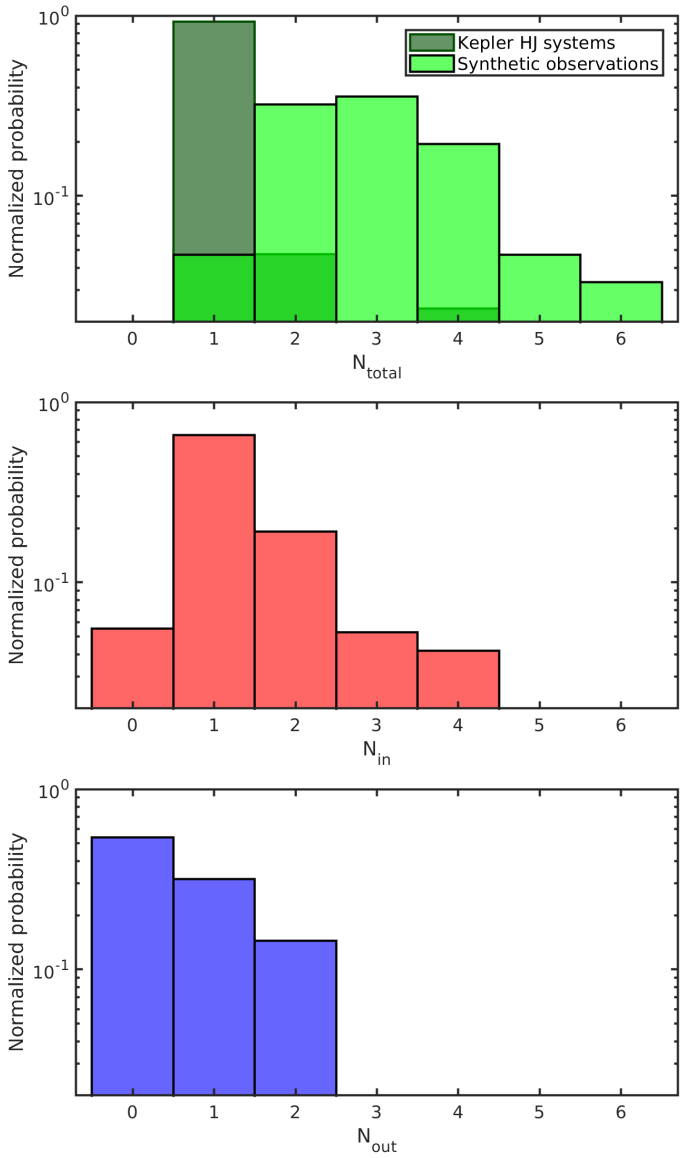

Figure 5.13.: Normalized probabilities of the final multiplicity, N, of the synthetically observed hot Jupiter systems arising from the equal-mass model. The subscript 'total' represents the total multiplicity of the systems (top panel). The bars coloured in dark green are the hot Jupiter systems observed by Kepler, and in light green are the synthetically detected hot Jupiter systems from the simulations. The value of $N_{\rm total} = 1$ corresponds to only the giant in the systems being detected. The subscripts 'in' (middle panel) and 'out' (bottom panel) represents the numbers of planet detected interior or exterior to the giant.

5.5. Impact of varying gas envelope accretion prescription

The gas envelope accretion routine is an important ingredient that affects the giant planet formation rate. In section 5.4, we adopted equation 5.12 to accrete gas envelopes during the phase when the planet is embedded in the disc, and this prescriptions is insensitive to local disc conditions as it is based on calculations performed at 5.2 au (see figure 5.2, dashed lines). Realistically, the local gas accretion rate is dependent on local disc properties, such as the local temperature and opacity. In this section, we investigate the impact of adopting equation 5.13, which depends on local disc conditions, on the planetary systems formed in the simulations.

5.5.1. Simulation outcomes

To investigate the changes to our main results (section 5.4) that arise when using a more realistic gas envelope accretion model, we ran an extra set of equal-mass simulations with equation 5.13. Equation 5.13 does not have a finite solution if $M_{\rm ge} = 0$ due to the negative power index, hence we initialise the protoplanets with a small and dynamically negligible envelope with $M_{\rm ge} = 10^{-8}$ $M_\oplus$ for all the protoplanets in the simulations. We consider three sets of runs with different values of the opacity reduction factor $f_{\rm opa} = 10^{-1}$, 10^{-2}, and 10^{-3}, and we use the labelling convention `o1`, `o2`, and `o3` when describing the simulations. Together with the initial K-values ($K = 4$ and 5) and the five selected systems, there are a total of 30 templates (3 sets of $f_{\rm opa}$ $\times$ 2 sets of initial K-value $\times$ 5 selected systems).

Equation 5.13 produces a larger gas accretion rate for a smaller value of $f_{\rm opa}$, as expected. We find that the `o1` and `o2` runs do not produce any giants with mass > 100 $M_\oplus$, while the `o3` runs have a giant formation rate of $\sim 6\%$. Hence, compared to the giant formation rate of the equal-mass runs in the main result (section 5.4.2), the `o3` runs produce more giants while the `o2` runs produce fewer. Synthetically observing the `o3` runs as described in section 5.4.4 yields a 1.36% chance that a detected system contains a transiting hot Jupiter and at least one transiting inner super-Earth.

Figure 5.14 and 5.15 show the dynamical evolution of a run from the template of `WASP47-XM-4K-o3` , where a hot Jupiter orbiting at ~ 0.06 au is formed with an inner super-Earth orbiting at ~ 0.04 au. The `o3` runs which successfully formed one hot Jupiter, follow a very similar evolution as our main set of simulations (e.g. figure 5.8) and always contain an inner planet.

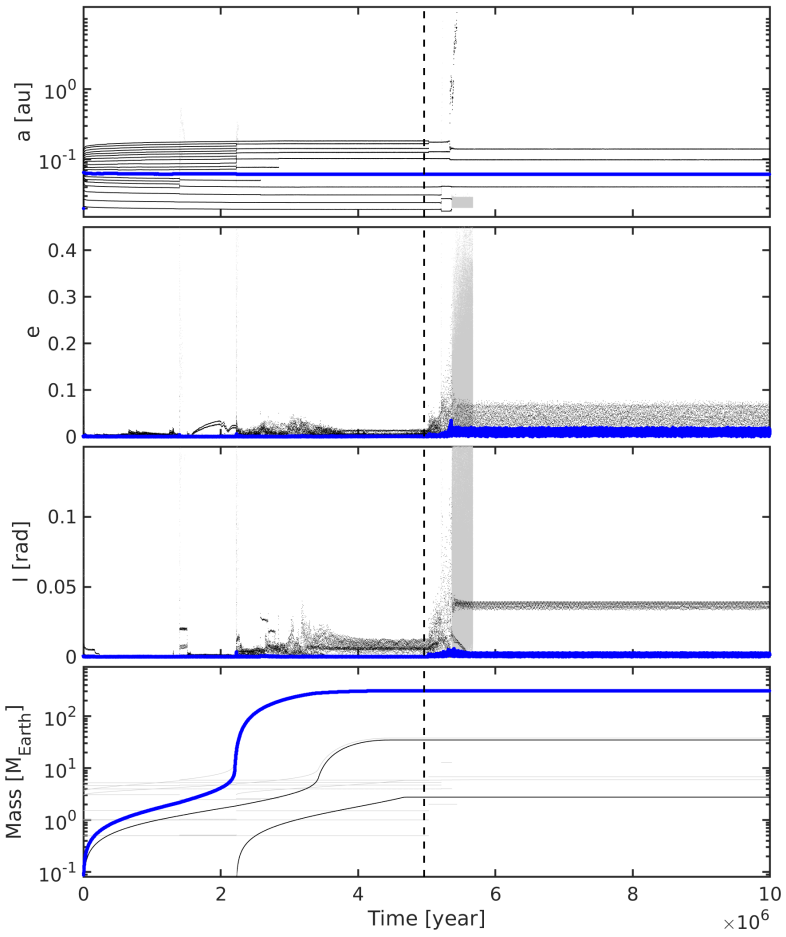

Figure 5.14.: Similar to figure 5.4 but for a run of the `WASP47-o3-4K` template, showing the evolution of the semimajor axes, eccentricities, inclinations, and masses. The blue lines indicated the most massive object in the simulation.

Interestingly, $\sim$ 3% of our `o3` runs produce two hot Jupiters in a system. Figure 5.16 and 5.17 show one example. A massive core ($\sim$ 5 $M_\oplus$) forms at an early stage at $\sim$ 0.1 Myr, and undergoes runaway gas accretion and opens a gap (figure 5.17, second panel). The rapid growth of a giant induces collisional accretion and formation of a second core that undergoes runway gas accretion followed by gap formation (figure 5.17, third panel). These two planets then grow at the viscous supply rate (equation 5.14) and end up as gas giants with masses $\sim$ 398 $M_\oplus$ and 542 $M_\oplus$.

The two giants in this run did not evolve much in terms of their semimajor axes. However, stronger dynamical interactions between the two giants result in higher

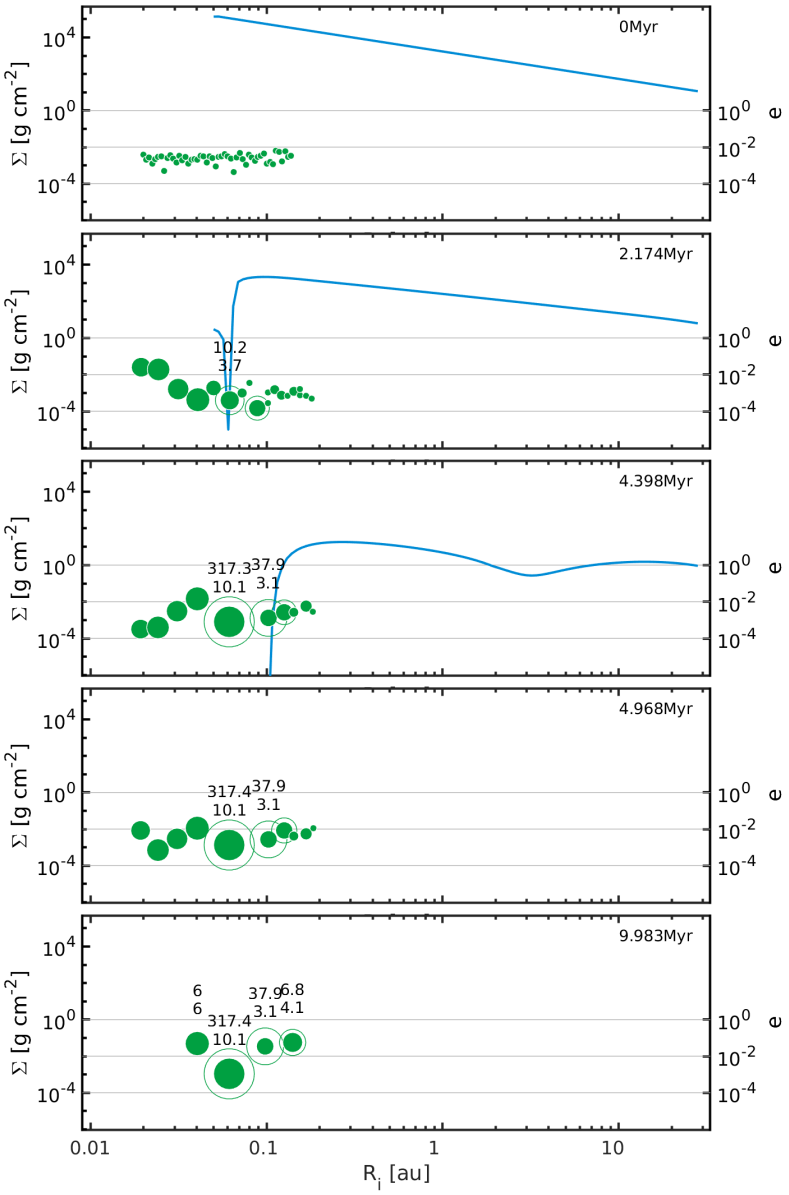

Figure 5.15.: Similar to figure 5.5 but for the same run as shown in figure 5.14, an equal-mass run with the new gas envelope accretion routine. As with the other equal-mass runs that form a giant planet, the gap opening, photoevaporation, and disc dissipation times are similar to the seed-model runs. The bottom panel shows a time close to the end of the simulation, and the system has a hot Jupiter at ~ 0.06 au and one inner super-Earth.

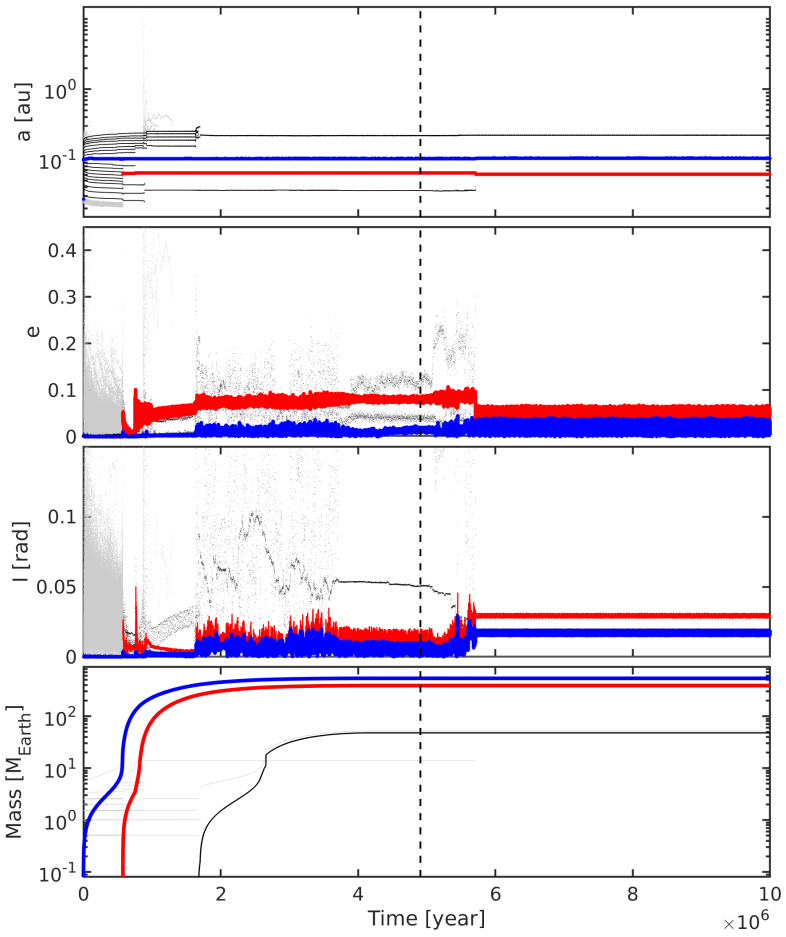

Figure 5.16.: Evolution of a run from the `Kepler730-o3-4K` template, similar to figure 5.4. The blue lines indicate the most massive planet the simulation, and the red lines indicated the second-most massive object.

eccentricities and inclinations (figure 5.16, blue and red lines) than for single hot Jupiter systems (e.g figure 5.4 and 5.14, blue lines). Recent observations of the WASP-148 system show that it may have two close-in giants orbiting and transiting the same star (Hébrard et al., 2020). The WASP-148 system contains two giants that both have mass ~ 100 $M_\oplus$ and semimajor axes ~ 0.08 and 0.21 au. The two planets WASP-148 b and c are on more eccentric orbits ($0.2 < e < 0.36$) than obtained in our simulation (see figure 5.16), and this may be because the two giants in our simulation did not experience strong scattering with another giant planet.

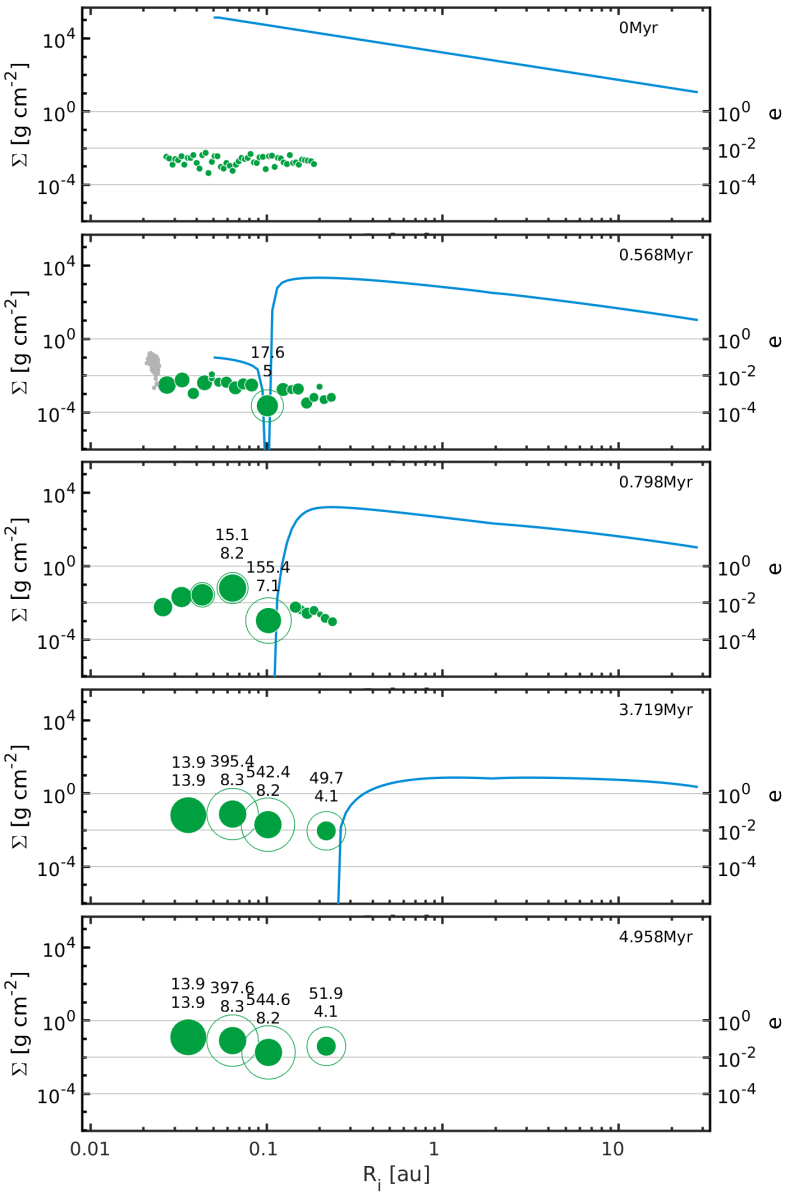

Figure 5.17.: Similar to figure 5.15 but for the same run as shown in figure 5.16. The first gap opening planet forms early ($\sim$ 0.5 Myr). Another large core forms before 1 Myr, allowing two gas giants to form before the gas disc dissipates.

5.6. Discussion and conclusions

In this study, we have presented the results of N-body simulations of the *in situ* formation of planetary systems containing a hot Jupiter and super-Earths that orbit interior to the giant planet. The aim is to examine whether or not planetary systems with this architecture can form locally by the collisional accretion of planetary embryos followed by gas envelope accretion onto cores that grow to sufficient mass.

This study is motivated by observations that demonstrate the existence of planetary systems with the above described architecture. In particular, the systems WASP-47, Kepler-730, TOI-1130, Kepler-487, and Kepler-89 all have a transiting giant with an orbital period less than 30 days and contain at least one inner transiting super-Earth/sub-Neptune. We use these systems as templates to construct the initial conditions for the N-body simulations, where the orbital locations of the giant planets define the median semi-major axes of the planetary embryos.

Two different sets of initial conditions were considered, a 'seed-model' and an 'equal-mass model'. In the seed-models, a relatively massive seed-protoplanet (4 or 4.5 $M_\oplus$) was placed at the reference semi-major axis, with multiple 0.5 $M_\oplus$ embryos orbiting interior and exterior to the seed. The equal-mass models have a chain of equal mass embryos (0.5 $M_\oplus$) centred at the reference semi-major axis. The purpose of the seed-model was to study the dynamical evolution of systems in which a giant planet was essentially guaranteed to form, while the equal-mass simulations investigate the giant planet occurrence rate using a set of unbiased initial conditions.

The N-body simulations included a realistic collision model and a protoplanetary disc modelled as an α-disc subject to photoevaporation. The disc provided eccentricity/damping forces on the protoplanets that were also able to accrete gas from the disc. Orbital migration through disc-planet interactions was neglected.

We observed that the dynamical evolution of the seed-models consistently followed 4 phases of evolution: 1) an early impact phase, where embryo-embryo collisions occur frequently at early times because of the initial compact configurations of the embryos, and where the systems stabilize after this initial epoch of collisional evolution due to the disc damping forces; 2) a runaway gas accretion phase, where the seed-protoplanet undergoes runaway gas envelope accretion, and the rapid increase in the mass of this planet dynamically heats up other bodies in the system; 3) an outer disc damping phase, where the embryos exterior to the seed protoplanet continue to experience the disc damping forces and the interior planets experience almost no disc damping force because of gap formation; 4) a late chaotic phase, where giant impacts between planets are common due to the dispersal of the gas

disc.

By design, the seed-model is efficient at producing systems of coexisting hot Jupiters and inner super-Earths/sub-Neptunes. The final average multiplicity of the inner systems is 2.2, while the average outer system multiplicity is 3.5. The inner system planets are more massive than the outer planets on average. The average inner system planet mass is 13.2 $M_\oplus$, and that of the outer system is 4.8 $M_\oplus$. All seed-model runs result in a final system consisting of a hot Jupiter and inner super-Earths/sub-Neptunes (except for the `WASP47` templates for reasons discussed in section 5.4.1).

The equal-mass models require efficient collisional accretion among the embryos to occur in order for a giant planet to form within the gas disc lifetime. Only $\sim 1\%$ of the equal-mass runs produced a gas giant planet. The formation history for such systems follows the four phases described above for the seed-model. For the runs where no giant forms, the disc damping forces are dominant throughout the gas disc lifetime, and embryo collisions are only common at the very early stages of the simulations (due to the initial compact configurations) and after the gas disc disperses.

For the equal-mass runs, we undertook synthetic transit observations of our final planetary systems, and the proportion of the detected systems that contained a hot Jupiter and at least one inner super-Earth was 0.26%, similar to the occurrence rate of $\sim$ 0.2% for such systems from actual transit surveys. However, we find that there is only a $\sim 5\%$ chance that a hot Jupiter would be detected as a single planet without nearby interior or exterior super-Earths, and so it is clear that the model presented here cannot explain the majority of hot Jupiter systems. The final planetary systems we form containing a hot Jupiter and nearby super-Earths are too flat to agree with the results of transit surveys such as Kepler.

A crucial ingredient in the models is the gas envelope accretion prescription, and we examined the impact of adopting different gas accretion routines. This included the simple fit from (Coleman and Nelson, 2016b) to the 1D gas accretion simulations conducted at 5.2 au by Movshovitz et al. (2010), and a new fit to a large suite of 1D gas accretion simulations conducted at different orbital radii and with different envelope opacities using the 1D envelope structure model of Coleman et al. (2017). One conclusion from our study is that a significant opacity reduction factor of, $f_{\rm opa}$, is required to form a hot Jupiter in these simulations. A value of $f_{\rm opa} = 10^{-2}$ is not small enough, whereas $f_{\rm opa} = 10^{-3}$ results in approximately 6% of the equal-mass simulations forming a hot Jupiter, which is higher than the $\sim$ 1% rate obtained when using the Coleman and Nelson (2016b) gas accretion prescription. Due to

the high efficiency of giant formation with $f_{\rm opa} = 10^{-3}$, $\sim 3\%$ of the runs formed systems with two giant planets, similar to the recently reported planetary system, WASP-148 (Hébrard et al., 2020).

The simulations presented here show that within the parameter space that we have considered, the formation of systems containing a hot Jupiter and at least one interior super-Earth/sub-Neptune can form *in situ* through collisional accretion in a compact chain of planetary embryos, followed by gas accretion onto a core that grows to be of sufficient mass to undergo runaway gas accretion. Overall, the frequency with which such systems are detected when we synthetically observe the simulation outcomes is similar to the frequency of occurrence in actual transit surveys. However, closer inspection of the distribution of system multiplicities arising from the simulations shows that they do not match the observations, as the models predict that hot Jupiters should be rarely be detected as single planets.

The model presented here can form systems similar to WASP-47, Kepler-730, and TOI-1130, and may represent the means by which systems with these particular architectures formed. The failure to form hot Jupiter systems that appear as single planets in transit surveys, however, suggests an alternative formation scenario is required to explain the majority of hot Jupiter systems.

6. Conclusions and future work

讓我對這世界好奇
讓我信自己的真理

— 陳奕迅, 林夕 ⟨我的快樂時代⟩

6.1. Conclusions

The research presented in this thesis focus on the simulation of formation and dynamical evolution of planetary systems, and statistical comparison between different planetary system formation scenarios and observational data. To achieve a comprehensive simulation model, I adopted a new version of N-body code (by implementing SYMBA), that includes a realistic collision algorithm (section 2.1) and protoplanetary disc evolution for planet-disc interaction (section 2.2). Moreover, I have written a synthetic transit observation program, which can provide synthetic transit surveys on planetary systems from simulations (section 2.3), to provide meaningful comparison to the transit surveys data. I have used these programs to study different planetary system formation scenarios, such as the *in situ* formation and self-scattering between protoplanets (chapter 3), dynamical heating of inner planetary system by cold giants (chapter 4), and the coexistence of hot Jupiter and inner super-Earth companions (chapter 5).

In situ self-scattering evolutions

I have studied the formation of compact systems of super-Earths via dynamical instabilities and giant impacts. Using N-body simulations with the implemented collision model, I examine the *in situ*, final gas-free stage assembly of multi-planet systems via the collisional accretion of protoplanets. My approach to creating initial conditions was to take a few known Kepler 5-planet systems, and to use these as templates for producing systems of 20 protoplanets, whose total mass was the same as the original Kepler systems, but in a more compact configuration. During the N-

body simulations, the protoplanets would experience dynamical instabilities which lead to planet-planet scatterings and giant impacts. Two different prescriptions (perfect merger and realistic accretion model) for treating planetary collisions were adopted.

Simulations of this scenario produce the mutual Hill separation distribution peak at 20 mutual Hill radii, which is in agreement with the Kepler observation. However, I found that the simulated systems overproduce, by about a factor of 2, the numbers of high-multiplicity versus low-multiplicity systems compared to the Kepler systems. My simulations and the synthetic observations also produce single-transiting-planets with systematically larger eccentricities than the planets in multi-transiting-systems, $\{\langle e_1 \rangle, \langle e_{\geq 2} \rangle\} \approx \{0.1, 0.05\}$, because the single-transiting-planets are from systems that have undergone stronger scattering than the multi-transiting-systems, but still cannot generate a large enough eccentricity distribution to provide a good match to the observationally inferred distributions of eccentricities for single-transiting-systems from Mills et al. (2019) where $\{\langle e_1 \rangle, \langle e_{\geq 2} \rangle\} \approx \{0.21, 0.05\}$. The observed distributions of planetary multiplicities or eccentricities are not reproduced, because scattering does not excite the systems sufficiently.

For the same reason, the systems I explored did not dynamically excite themselves sufficiently for collisions to be highly disruptive, the simulation outcomes had almost no detectable dependence on the collision model adopted. Nevertheless, the realistic accretion model is important during post-processing the collision outcomes. It suggests that they would not significantly change the ice fractions of initially ice-rich protoplanets, but significant stripping of gaseous envelopes appears likely. Hence, it may be difficult to reconcile the observation that many low-mass Kepler planets have H/He envelopes with an *in situ* formation scenario that involves giant impacts after dispersal of the gas disc. In addition, I find that co-orbital planet pairs in horseshoe or tadpole orbits can be formed under this scenario, where around 1% of the simulations can produce a stable 1:1 resonance planet pair.

Dynamical heating by cold giants

In this project, I have studied the origin of the eccentricity dichotomy displayed by compact super-Earths which experienced dynamical heating by cold giants. Recent Kepler analyses have reported a dichotomy in the eccentricity distribution displayed by systems where a single planet transits compared with that displayed by the multi-planet systems (Xie et al., 2016; Mills et al., 2019). As mentioned in the last section, the *in situ* self-scattering simulations can provide results that are generally

in agreement with the data on the mutual separations between the planets. But when planetary system masses are adopted that are characteristic of those inferred for the Kepler multi-systems, then the degree of gravitational scattering is insufficient to provide large enough mutual inclinations or eccentricities to explain the multiplicity and eccentricity dichotomies (chapter 3 and Poon et al., 2020). This has led to an alternative hypothesis for explaining the eccentricity dichotomy, namely that giant planets in the outer regions of planetary systems perturb the inner systems.

Using N-body simulations, I showed that the dichotomy can arise if inner systems of super-Earths are frequently accompanied by outer systems of giant planets that can become dynamically unstable and perturb the inner systems. The initial conditions are constructed using a subset of the known Kepler 5-planet systems as templates for the inner systems, and systems of outer giant planets with masses between those of Neptune and Saturn that are centred on orbital radii between 2 to 10 au. The parameters of the outer systems are chosen so that they are always below an assumed radial velocity detection threshold of ~ 3 ms^{-1}.

I found that the synthetic transit observation of the simulated systems considered produce better agreement with the Kepler data compared to the multiplicity ratios obtained from the self-scattering of the inner systems of super-Earths. Some of the outer templates resulted in inner systems that are in very good agreement with the Kepler multiplicity ratios and reproduce the Kepler multiplicity dichotomy. Synthetic observation of the simulated inner systems also produces a very clear eccentricity dichotomy. For the simulations with outer systems, I obtain the mean eccentricities $0.06 \leq \langle e_{\geq 2} \rangle \leq 0.10$ for multi-transiting systems and $0.11 \leq \langle e_1 \rangle \leq 0.25$ for single-transiting systems. The outer system comprised of 6 planets centred around a semi-major axis of 5 au with masses of 60 $M_\oplus$ resulted in an eccentricity dichotomy characterised by $\{\langle e_1 \rangle, \langle e_{\geq 2} \rangle\} \approx \{0.25, 0.07\}$, which is in decent agreement with the values $\{\langle e_1 \rangle, \langle e_{\geq 2} \rangle\} \approx \{0.3, 0.04\}$ reported by Xie et al. (2016) and the value $\{\langle e_1 \rangle, \langle e_{\geq 2} \rangle\} \approx \{0.21, 0.05\}$ reported by Mills et al. (2019).

Hot Jupiters with companion super-Earths

Recent observations confirmed multiple systems where a hot Jupiter and inner super-Earths coexist, such as the WASP-47, Kepler-730, and TOI-1130 systems. The origin and the formation process for such systems are not yet well understood. I examined the hypothesis of forming this type of system in a protoplanetary disc while assuming an *in situ* formation process by using N-body simulations, as well as the adopted planet-disc interaction and the realistic collision algorithm. In the *in situ* disc, the

migration forces are switched off, while other physical process remain active, such as eccentricity and inclination damping, disc gap opening, and especially, protoplanet gas envelope accretion. With the gas envelope accretion routine, protoplanets can accrete their gas envelope and eventually become gas giants if they can accrete gas quick enough before the disc is dissipated.

I constructed the initial conditions by using the known positions of the hot Jupiter in such systems as the reference locations (e.g. semi-major axis of WASP-47b at 0.051 au, Kepler-730b at 0.069 au, TOI-1130c at 0.071 au). Two different sets of initial conditions are considered, a so-called 'seed-model' and a 'equal-mass-model'.

The seed-model is constructed by putting a relatively massive protoplanet ($\sim$ 4 $M_\oplus$) at the reference semi-major axis together with multiple 0.5 $M_\oplus$ protoplanets sitting both closer and further to the side of the reference protoplanet. The focus of the seed-model is to study the dynamical evolution of a system with a growing giant. Simulations of the seed-model lead to the formation of systems contaning a hot Jupiter and nearby super-Earths/sub-Neptunes. The evolution consistently follows four distinct phases: early giant impacts; runaway gas accretion onto the seed protoplanet; disc damping-dominated evolution of the embryos orbiting exterior to the giant; a late chaotic phase after the dispersal of the gas disc.

The equal-mass-model is constructed by putting a chain of equal mass protoplanets (0.5 $M_\oplus$) centred at the reference semi-major axis. In the equal-mass-model, giant planet formation requires protoplanets to gain mass by accreting other protoplanets in the early stage of the simulation, which can lead to the gas envelope accretion routine to be significant enough to form giants. Different formation parameters have also been tested in this equal-mass-model, such as the different values of initial mutual Hill separation ($4 \leq K \leq 5$) and different values of opacity reduction factor (which affect the gas accretion rate of the protoplanet). Approximately 1% of the equal-mass simulations form a gas giant planet and follow the same four-phase evolution. Synthetic transit observations of the equal-mass simulations provide an occurrence rate of 0.26% for the systems containing a hot Jupiter and an inner super-Earth, similar to the 0.2% occurrence rate from actual transit surveys. Also, a small subset of my simulations form two close-in giants, similar to the WASP-148 system. However, the simulated hot Jupiters are almost always detected in multi-transiting systems, in disagreement with observations.

Concluding remarks

How did exoplanetary systems form and evolve to create the diversity that we now observe? The few studies presented in this thesis are set to address this question. We can see from the results that the *in situ* scenario considered throughout this thesis provided some outcomes that are comparable to the observations. However, there are also important observational behaviours that are yet to be explained by the *in situ* model. Compared to other existing models, e.g. migration, the *in situ* model has both advantages and disadvantages in connecting the observations and theoretical studies. It was never the intention of this thesis to rule out the plausibility of either the *in situ* or migration model, but I hope that the results presented in this thesis have been encouraging enough, to some extent, to merit further investigation of the *in situ* model.

6.2. Future work

N-body simulations are a very powerful tool to provide studies of planetary system evolutions, but there is still no single model which can reproduce all the observational statistic. Future transit missions, such as PLATO (Rauer et al., 2014), will monitor many more stars than Kepler, together with the high precision RV observations, such as the Thirty Meter Telescope (TMT, Skidmore et al., 2015), they will provide a much larger and more complete set of statistical data and place more stringent constraints on the formation histories of exoplanetary systems. With this large and continuously growing number of confirmed exoplanets, the study of the comparison between planetary system formation theories and exoplanet observational statistics will play an important role in the future. In the past decade, many formation studies were focused on one specific type of observed system such as hot Jupiters in single-planet systems, transiting close-in and compact systems of super-Earths, or cold giants discovered by radial velocity. But recently, more diverse exoplanet systems were observed, such as, WASP-47 which is a system where a hot Jupiter and inner super-Earths coexist, and WASP-148 which contain 2 inner giants. The formation history of such systems is not yet well understood.

To improve the understanding of the formation history, two approaches on improving the N-body simulation model can be taken, which are 1) develop a global simulation model which includes higher order physical effects in the simulation to provide more in depth and detailed evolution information, and 2) with the existing model, apply a significantly broader range of initial conditions. With these two ap-

proaches, I would like to propose the following studies which will improve the planet formation theory and the numerical simulation model.

Tracking the compositional change of planets during giant impacts

The components of a planet can be classified into three groups: rocky core, ice/water envelope, and gas envelope. In a traditional N-body simulation, the planets are only considered in the form of uniform density, instead of separating them into three different components. This assumption can be made, up to the level of assuming perfect merger during a planet-planet collision, where the collision outcome retains all the masses from the two colliders. In a more realistic collision event, while the impact energy and the impact angle are high enough to generate mass losses in the form of fragmentations, the surface gas component is easiest to get stripped off, then it is the ice/water component, and this is followed by the rocky component.

Planet-planet collision outcomes can be very diverse. The collision outcome is highly depending on the collision condition, such as the impact energy, impact angle, and the composition of the colliding objects. High-resolution smoothed-particle hydrodynamics (SPH) simulations of planetary collisions can provide systematic results and prescriptions for the collision outcomes. Different parameters can be applied to the SPH experiments, such as the type of planet in terms of composition, the ratio between each composition, impact energy, and impact angle. An in-depth study of the high-resolution SPH results will be able to provide large improvement to the collision results that could be adopted by N-body simulations, such as the position, mass, and velocity distributions of the collision debris, and final masses and composition ratios of the remnants.

A prescription of planet temperature evolution can also be implemented to my N-body simulation model. The temperature evolution can include the increase of core temperature during giant impact events, where a portion of impact kinetic energy is converted to the form of thermal energy. The increase of core temperature would heat up the atmosphere, if the planet originally has one, which would expand the gas, decrease the Bondi radius of the planet, and lead to a potential mass loss for the gas envelope component. The cooling of the atmosphere and the planet core after the giant impacts can also be implemented to the temperature evolution model to allow the increase of Bondi radius, and re-accretion of the gas component. The prescriptions mentioned here would lead to changes in planetary radii and masses. These changes can result in different outcomes of planet-planet scattering and the realistic collision algorithm. Future observation by TMT may be able to detect

young, molten planets that have recently formed via giant impacts, and hence will provide constraints for models that involve evolution via giant impacts.

Together with the aid of my version of an N-body integrator which considered realistic collisions with fragmentations, tracking all these compositional changes throughout the whole N-body simulation, even before the disc dissipated, is possible. In a dynamically hot system, giant impacts can result in significant mass loss (e.g. coexistence of hot Jupiter and close-in compact super-Earths studied in chapter 5 and Poon et al. submitted.). Studying the dynamical evolution for such fragments can provide answers to some questions, such as: Would the ice/water fragments accrete back to the original planet, accrete by other planets, or even get scattered out the system? Can the atmosphere survive the giant impacts throughout the whole dynamical evolution process? How does this compositional change take part in shaping the final simulated system? Is the final simulation outcome comparable to observational data, such as the position of radius-valley or ratio of planet that has thick atmosphere?

Evolution of the dynamically unstable outer planetary systems

Previous studies showed that unstable outer planetary systems with Jovian size giants at a few au can induce perturbations to the inner systems (e.g. Mustill et al., 2017; Huang et al., 2017). On top of that, my recent study shows that the general scenario of outer unseen sub-Jovian planets (RV precision $\lesssim 3$ ms^{-1}) perturbing inner planetary systems is a promising mechanism for explaining some aspects of the Kepler multiplicity and eccentricity dichotomies (chapter 4 and Poon and Nelson, 2020). Such unseen outer planets are very difficult to detect with current instruments such as SPHERE on the VLT, GPI on Gemini and SCExAO on Subaru. Future RV observations and direct detection with high precision, such as TMT and its installed ExAO systems, will be able to unfold the planets population which is similar to some of the outer planetary parameters that I considered in my previous work, and demonstrate whether or not these systems exist with the frequency required to explain the Kepler data. Also, the ALMA detection of kinematic signatures of planets in molecular line observations (Pinte et al., 2019) points to a population of outer planetary systems. It is of interest to investigate how these systems might evolve towards becoming dynamically unstable by considering the inward migration during the disc phase.

I plan to run hydrodynamical simulations of the planetary disc, with the FARGO code, that explore this scenario. Through these simulations, we can determine under

what conditions the systems can be developed to the long-term unstable systems that can perturb the inner planetary systems. Constructing observables of these simulation results can be used to compare with the ALMA observations, which can provide a good connection between the scenario of outer giant planets perturbation to their inner systems and the observations.

Long term tidal evolution for close-in planets

A recent study by Owen and Campos Estrada (2020) suggested that a large number of short period super-Earths have lost their gas envelopes by atmosphere photoevaporation to explain the origin of the radius-valley, which suggested that many Kepler super-Earths would have a tidal dissipation parameter, $Q \sim 100$, because of their rocky composition. With this Q-value, the tidal damping time-scale for many close-in super-Earths would be shorter than 1 Gyr and result in very circular orbits. However, applying this Q-value would be difficult to reproduce the observed eccentricity dichotomy (chapter 4 and Poon and Nelson, 2020) where the mean eccentricity of Kepler single-planet system is $\langle e_1 \rangle = 0.3$ (Xie et al., 2016) or 0.21 (Mills et al., 2019).

Moreover, the semi-major axis distribution of Kepler single-planet system peaks at $\sim$ 0.05 au where tidal eccentricity damping can be significant. On the other hand, if the atmosphere can be retained for such close-in super-Earths, the value of Q could be increased to $\sim$ 1000. Which would provide, in general, a longer tidal evolution time-scale and allow the eccentricity dichotomy to be retained throughout the dynamical evolution of the planetary systems. Obtaining a distribution of the tidal dissipation parameter Q, which can yield a fit on both the observations and dynamical simulations, would provide information for a more complete formation model.

Some previous N-body studies on Kepler-like systems with close-in planets (semi-major axis $<$ 0.05 au) show that higher order physical effects with long evolution time-scale, such as tidal eccentricity damping, are not going to affect the statistical outcome significantly for simulations run up to a few 10 Myr time-scale. It is because a reasonable N-body simulation time-scale of such Kepler-like systems is in orders of 10 Myr due to the small time-step. However, the age of Kepler systems is in Gyr time-scale, and such effects with long evolution time-scale might become important while comparing the simulation to the Kepler observation statistic.

My plan is to test the relation between different distributions of Q-values and the eccentricity dichotomy signal from the synthetic observation of the N-body

simulation results. I will apply an adopted version of tidal evolution model on my N-body simulations, which can model the tidal evolution of planets in Gyr time-scale, and run simultaneously to the N-body simulations for 10 Myr time-scale. The adopted simulation can be used for other future N-body studies, which contain close-in super-Earths or sub-Neptunes, to provide long term tidal effect in a reasonable simulation time. And the result of the distribution of Q-value can provide hints on the underlying eccentricity distribution of Kepler observations, such as the population of low and high eccentricity systems.

A. Supplementary contents

A.1. Planetesimal ring fragmentation

For the imperfect collision simulations, which adopted the collision model from Leinhardt and Stewart (2012), approximately 10% of our simulations experienced at least one super-catastrophic collision, leading to the formation of a ring composed of collision debris in the form of planetesimals. This ring was often confined to the inner most regions close to the star, where collision velocities can reach their highest values, and this led to large numbers of particles needing to be integrated using small time step sizes. In order to overcome this problem, we developed a scheme for reducing the masses of the ring particles on a time scale corresponding to their collision time scales. Then, once the mass of the ring reaches negligible values, the ring particles could be removed from the simulations, since the planetesimals would then be ground down to dust, which would be removed in reality by the Poynting-Robertson effect for grain sizes between 1 mm to 1 μm (Poynting, 1903; Robertson, 1937; Guess, 1962) or radiation pressure for grains smaller than 1 μm (Burns et al., 1979). When there is a fragmentation ring detected between the inner-most big body and the host star, we apply the following step during the simulation.

The total number of bodies in the ring, N, the mean semi-major axis, $\langle a \rangle$, eccentricity, $\langle e \rangle$, inclination, $\langle I \rangle$, and semi-major axis of the inner- and outer-most bodies in the ring (a_{out} and a_{in}) can be found from the simulations. These can be used to calculate the collision time in the ring, $\tau_{\rm coll}$, according to

$$\tau_{\rm coll} = \frac{1}{n\sigma \langle v \rangle}, \tag{A.1}$$

where n is the number density given by,

$$n \approx \frac{N}{2\pi(a_{out}^2 - a_{in}^2)\langle a \rangle \langle I \rangle}. \tag{A.2}$$

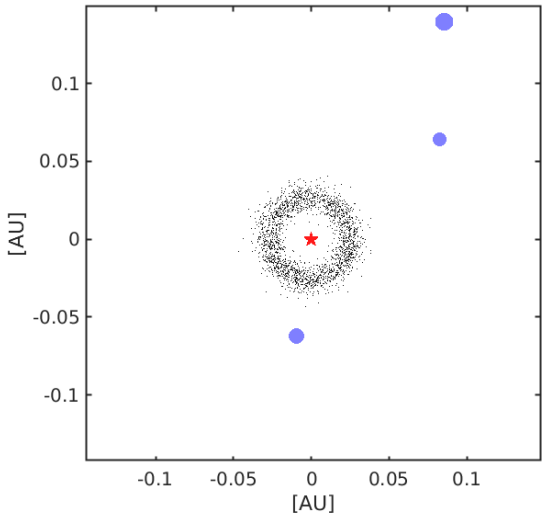

Figure A.1.: An example of an inner ring from one of our simulations, where the small black dots are the position of the inner ring objects. The blue circles denote the planets with their size is relative in term of their mass. The centre red pentagram is the position of the host star. Viewed from the top of the system all bodies are orbiting the host star in an anti-clockwise direction. Parameter information about this inner ring is listed in figure A.2 and table A.1.

The velocity dispersion $\langle v \rangle$ is given by,

$$\langle v \rangle \approx \langle v_k \rangle \sqrt{\langle e^2 \rangle + \langle I^2 \rangle}, \tag{A.3}$$

where $\langle v_k \rangle$ is the mean Keplerian velocity of the objects in the ring. The collision cross-section is simply,

$$\sigma = \pi R_p^2. \tag{A.4}$$

Having obtained the collision time, we then decrease the mass of the ring particles according to

$$m_p(t) = m_p(t_0) e^{-(t-t_0)/\tau_{\rm coll}}, \tag{A.5}$$

where t_0 is the time of ring formation, and after 15 e-folding times we remove the ring particles, since their masses are then negligible.

Figure A.1 shows the ring formed in one of our simulations. Figure A.2 shows the parameters of the ring shown in figure A.1, and table A.1 lists the ring parameters. We can see 2177 bodies with masses ~ 0.0015 $M_\oplus$ concentrated between 0.01 and 0.05 au. Applying the parameters of the ring listed in table A.1 to equations A.1, we get $\tau_{\rm coll} = 2.12$ yrs, leading to rapid removal of the ring.

We note Mustill et al. (2018) also include an imperfect accretion model in their

Table A.1.: Fragmentation ring parameters for `Kepler55_low_06`, displayed in figure A.1.

N	a_{in} [au]	a_{out} [au]	$\langle a \rangle$ [au]	$\langle e \rangle$	$\langle I \rangle$ [rad]
2177	0.0241	0.0283	0.0260	0.2213	0.1272

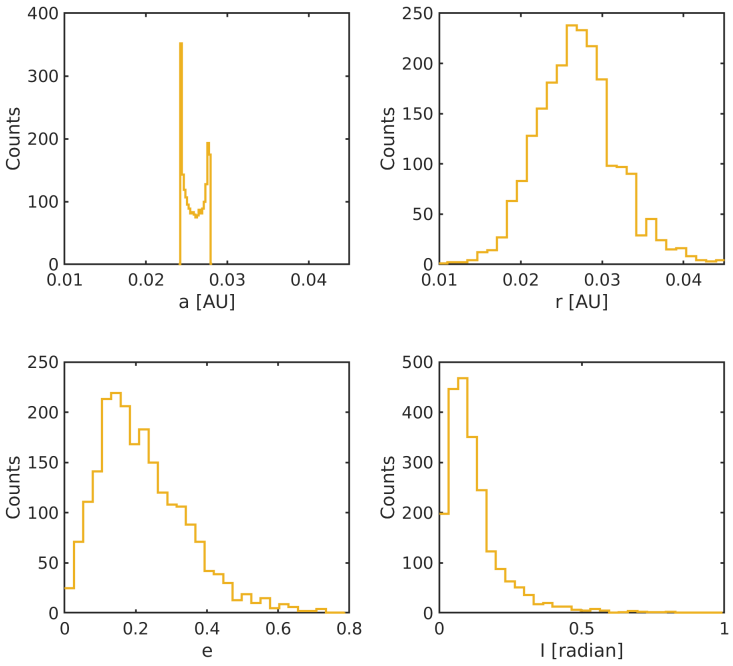

Figure A.2.: Counting distribution of all inner ring object parameters. Where it shows the distribution of the (top left) semi-major axis, (top right) distance, r, from the host star, (bottom left) eccentricities, and (bottom right) inclinations.

simulations, and immediately remove all small debris particles after they are formed because of the small time scale for collisions and collisional grinding versus the re-accretion time scale of the debris onto nearby protoplanets. We have examined the re-accretion time scale for debris particles in our simulations, and find that after a debris cloud is generated the time scale for half of the debris particles generated to be re-accreted is between $\sim$ 2 years up to $\sim$ 100 years, as shown in Figure A.3. This is considerably shorter than the re-accretion time scale of 30,000 years estimated by Mustill et al. (2018), but perhaps comparable to or longer than the typical grinding time scales for debris clouds generated by during collisions. This suggests that a realistic model for the evolution of post-collision debris should allow a fraction of it to

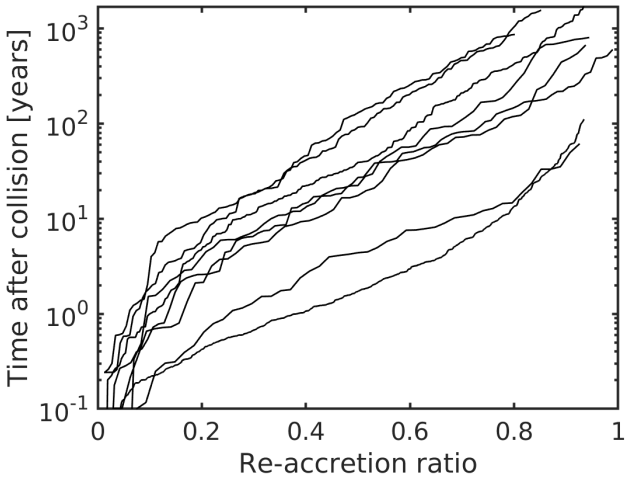

Figure A.3.: Time since collision versus the fraction of debris particles remaining since the collision for 8 collision events selected from the simulations.

re-accrete while the other fraction is ground down and removed by radiation pressure. Incorporating such a model, however, goes beyond the scope of this chapter.

A.2. Surface density fitting model

Table A.2 shows all the coefficients adopted to fit the surface density, Σ_{fit}, together with the type of fitting model chosen from the models described in section m3.2.3. The choice of fitting model was made by selecting the one that gave the best least-squares fit.

Table A.2.: Surface density fitting model for the eight Kepler templates. For model details see equation 3.14.

System template	Fitting model	c_1	c_2	c_3	c_4
`Kepler55`	Power law	1.047×10^{-1}	-2.488	73.80	/
`Kepler80`	Polynomial	-1.687×10^{7}	2.597×10^{6}	-1.097×10^{5}	1.917×10^{3}
`Kepler84`	Polynomial	2.062×10^{5}	-9.379×10^{4}	1.142×10^{4}	-1.507×10^{2}
`Kepler102`	Fourier series	1.042×10^{2}	57.58	37.81	60.55
`Kepler154`	Polynomial	6.073×10^{4}	-3.512×10^{4}	5.303×10^{3}	-25.78
`Kepler169`	Exponential	5.233×10^{2}	-17.72	0	/
`Kepler292`	Power law	29.21	9.643×10^{-1}	0	/
`Kepler296`	Power law	9.924×10^{-1}	-2.082	0	/

A.3. Simulation outputs for individual templates

This section provides the simulation outcomes from all system templates that were not shown in the main body of chapter 3.

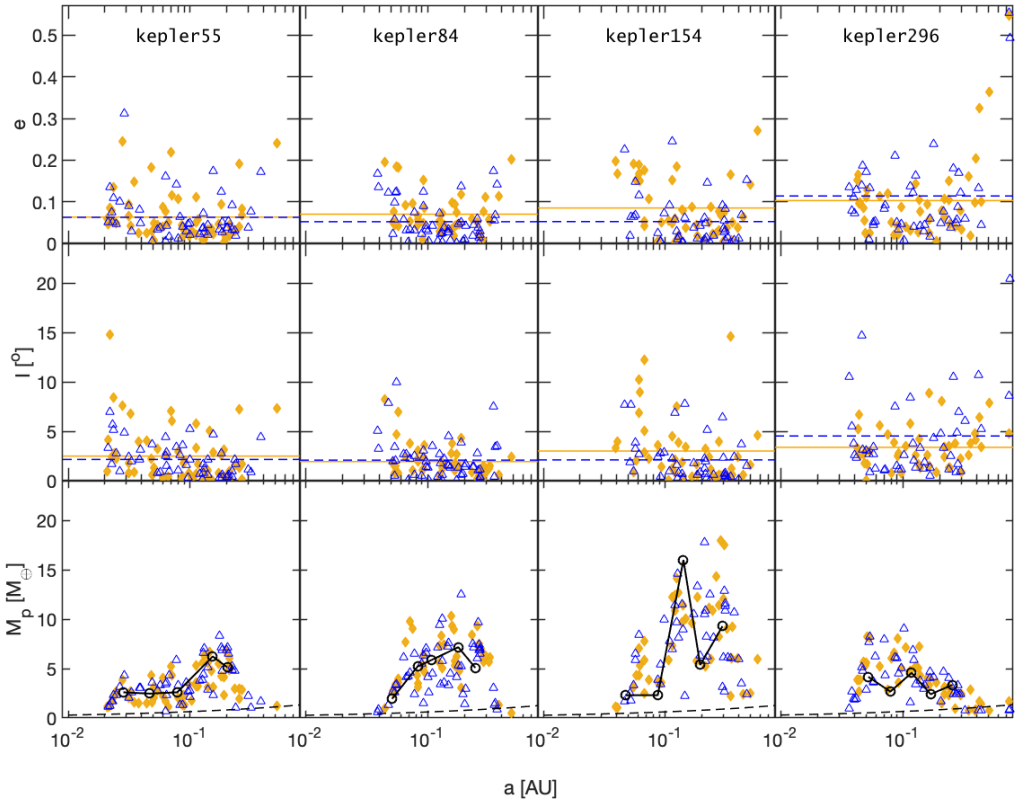

Figure A.4.: Same plot as figure 3.7 but for the imperfect Kepler-55, 84, 154, and 296 templates.

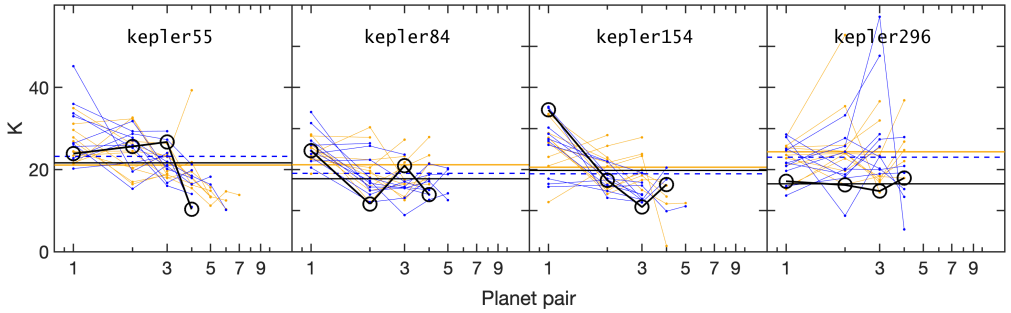

Figure A.5.: Same plot as figure 3.12 but for the imperfect `Kepler55`, `84`, `154`, and `296` templates.

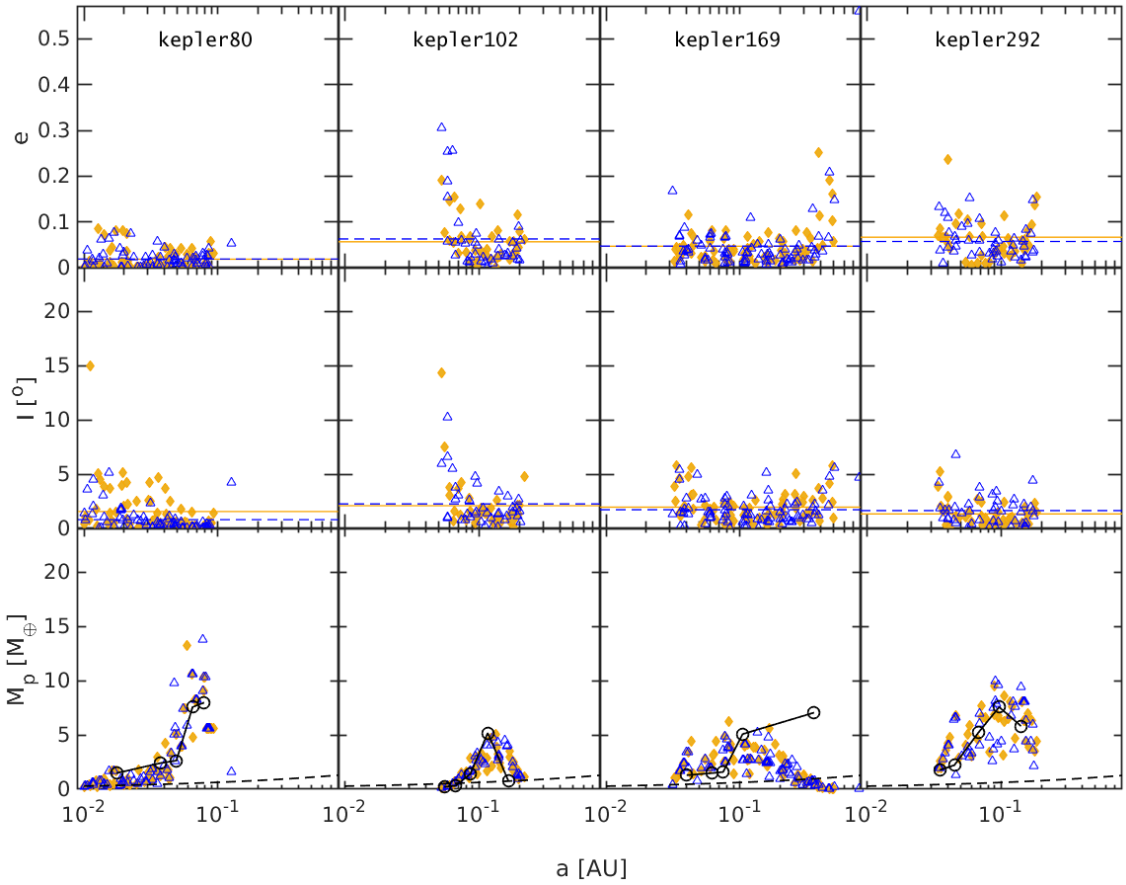

Figure A.6.: Same plot as figure 3.7 but for the perfect `Kepler80`, `102`, `169`, and `292` templates.

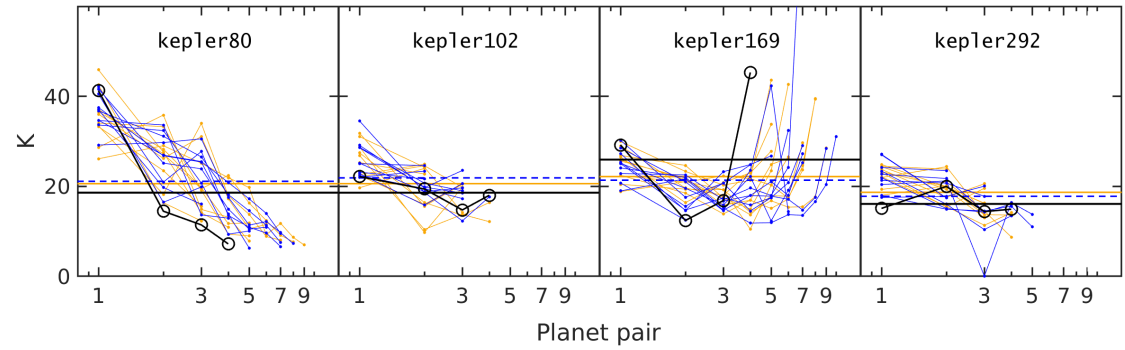

Figure A.7.: Same plot as figure 3.12 but for the perfect `Kepler80`, `102`, `169`, and `292` templates.

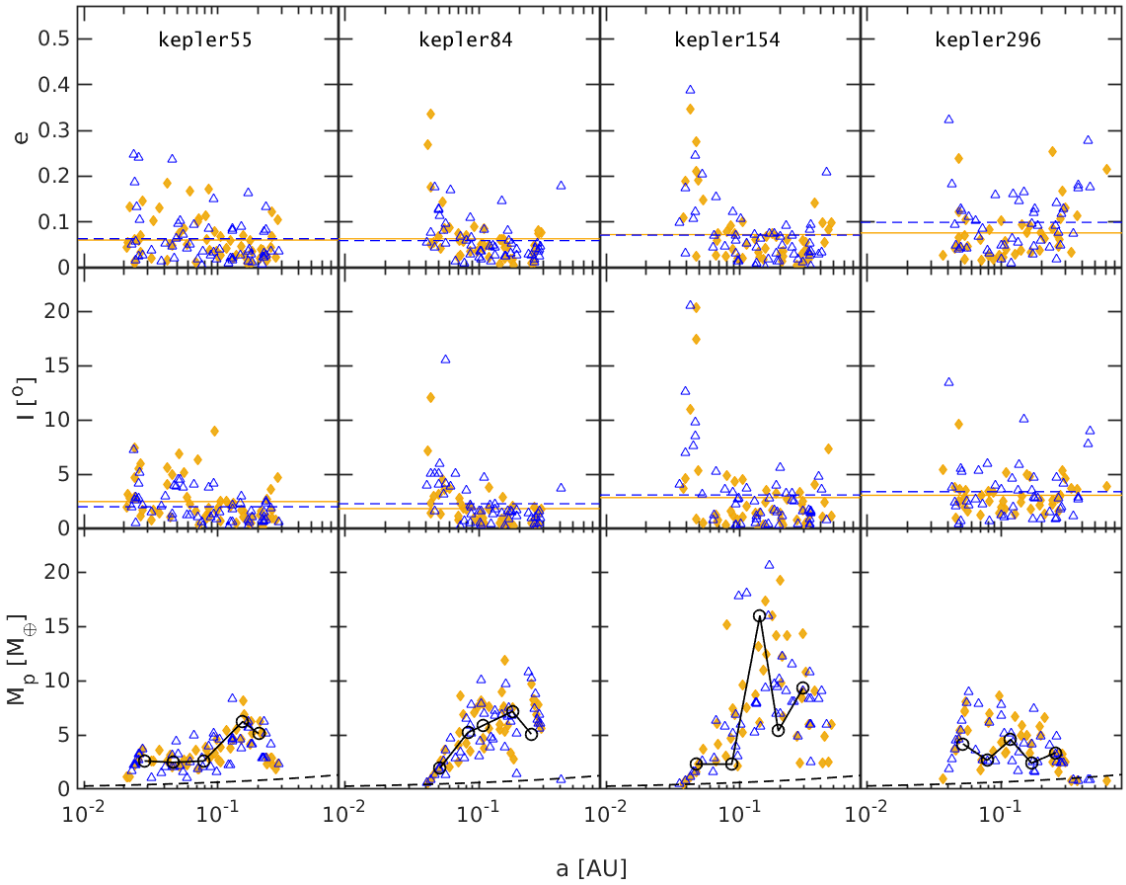

Figure A.8.: Same plot as figure 3.7 but for the perfect `Kepler55`, `84`, `154`, and `296` templates.

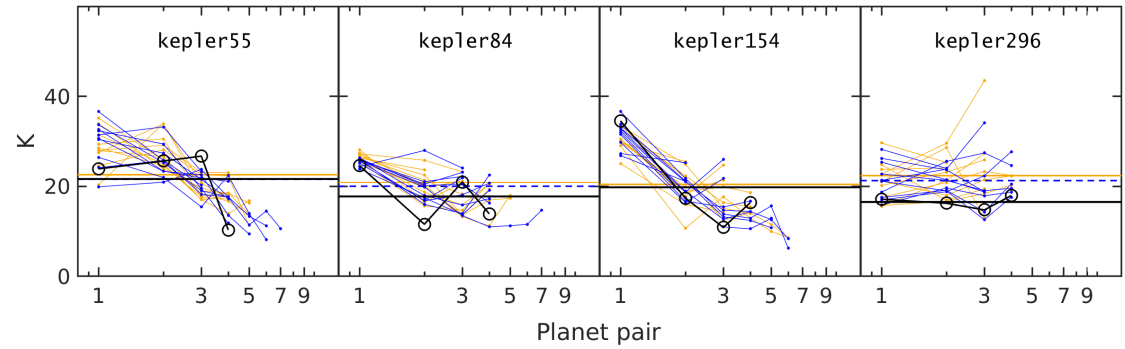

Figure A.9.: Same plot as figure 3.12 but for the perfect `Kepler55`, `84`, `154`, and `296` templates.

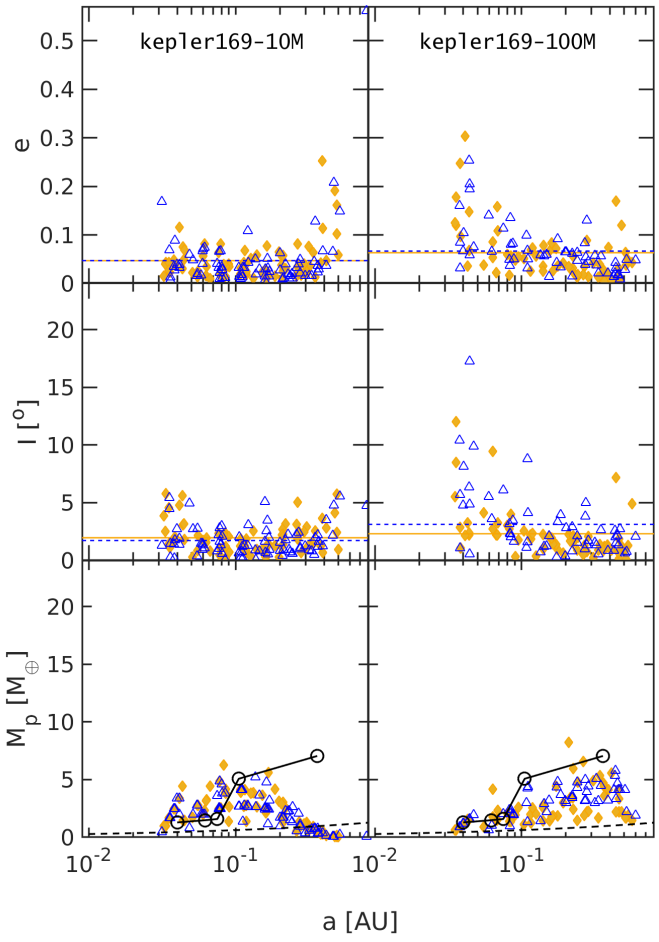

Figure A.10.: Similar to figure 3.7 but for the comparison of template `Kepler169` (perfect routine) after running 10 Myr and 100 Myr.

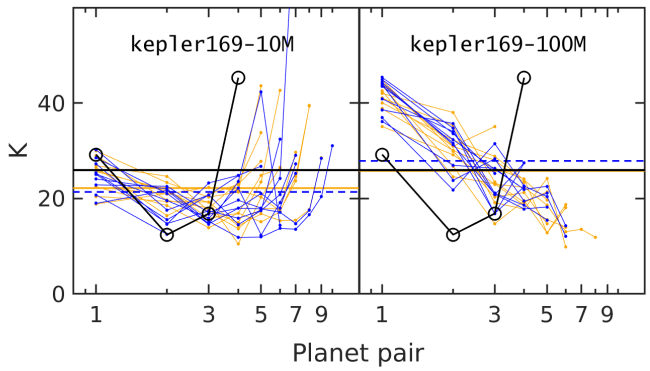

Figure A.11.: Similar to figure 3.11 but for the comparison of template `Kepler169` (perfect routine) after running 10 Myr and 100 Myr.

A.4. Initial eccentricities of the inner systems

The upper panel of figure A.12 shows the initial eccentricities for all planets contained in the inner systems considered in this study. These are the values that arise when setting the eccentricities according to a Rayleigh distribution, as described in Section 4.3.3. The lower panel of figure A.12 shows the final eccentricities after the inner planet systems have been evolved for 10 Myr in the absence of the outer systems of giant planets, and we see that some planets have experienced eccentricity growth even though scattering between the planets has been minimal. Figure A.13 shows the same final eccentricities after the inner planet systems have been evolved for 10 Myr in the absence of the outer systems of giant planets, but now on a system by system basis. In some systems there has been a redistribution of the angular momentum deficit (AMD) causing some of the innermost and lowest mass planets to experience increases in their eccentricities.

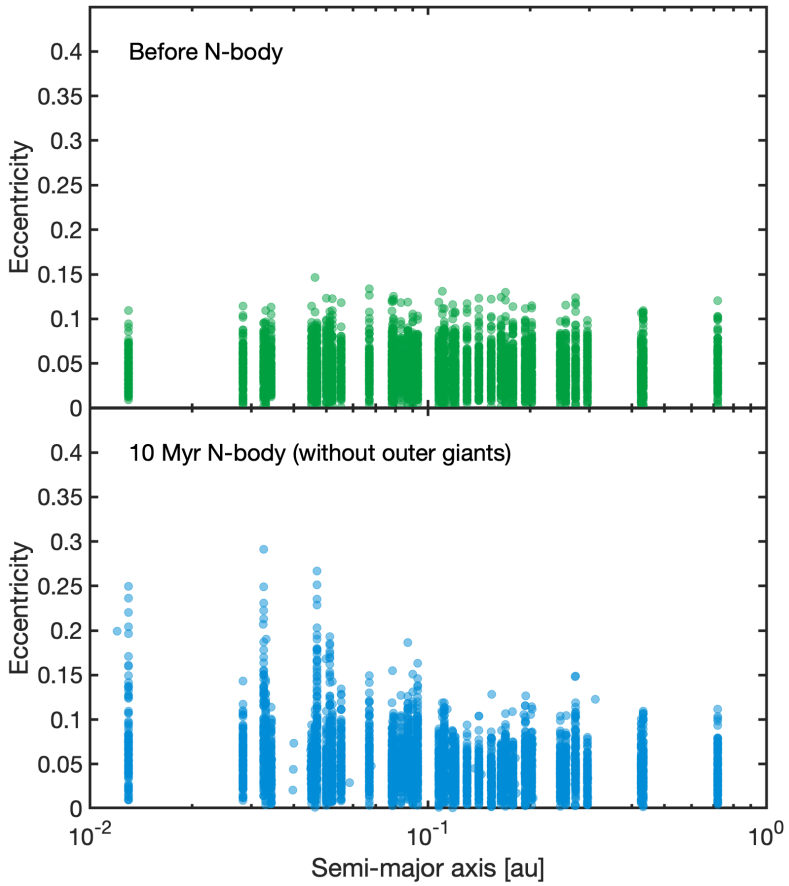

Figure A.12.: Eccentricities versus semi-major axis for the inner system before and after they have been evolved for 10 Myr in the absence of giant planets

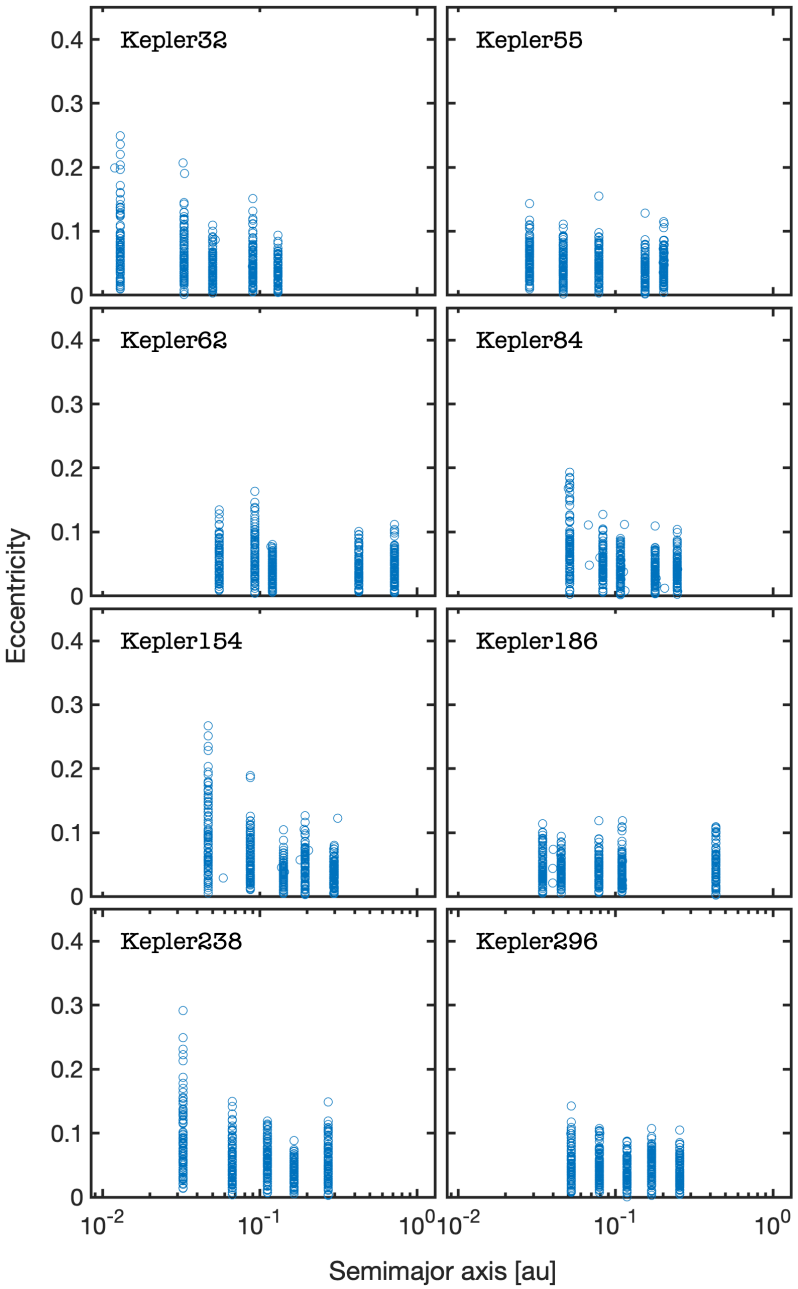

Figure A.13.: Eccentricities versus semi-major axis for the inner systems after they have been evolved for 10 Myr in the absence of giant planets

A.5. Eccentricity distributions obtained by synthetic observations

Eccentricity distributions of inner systems from 8 outer system templates obtained by synthetic observations.

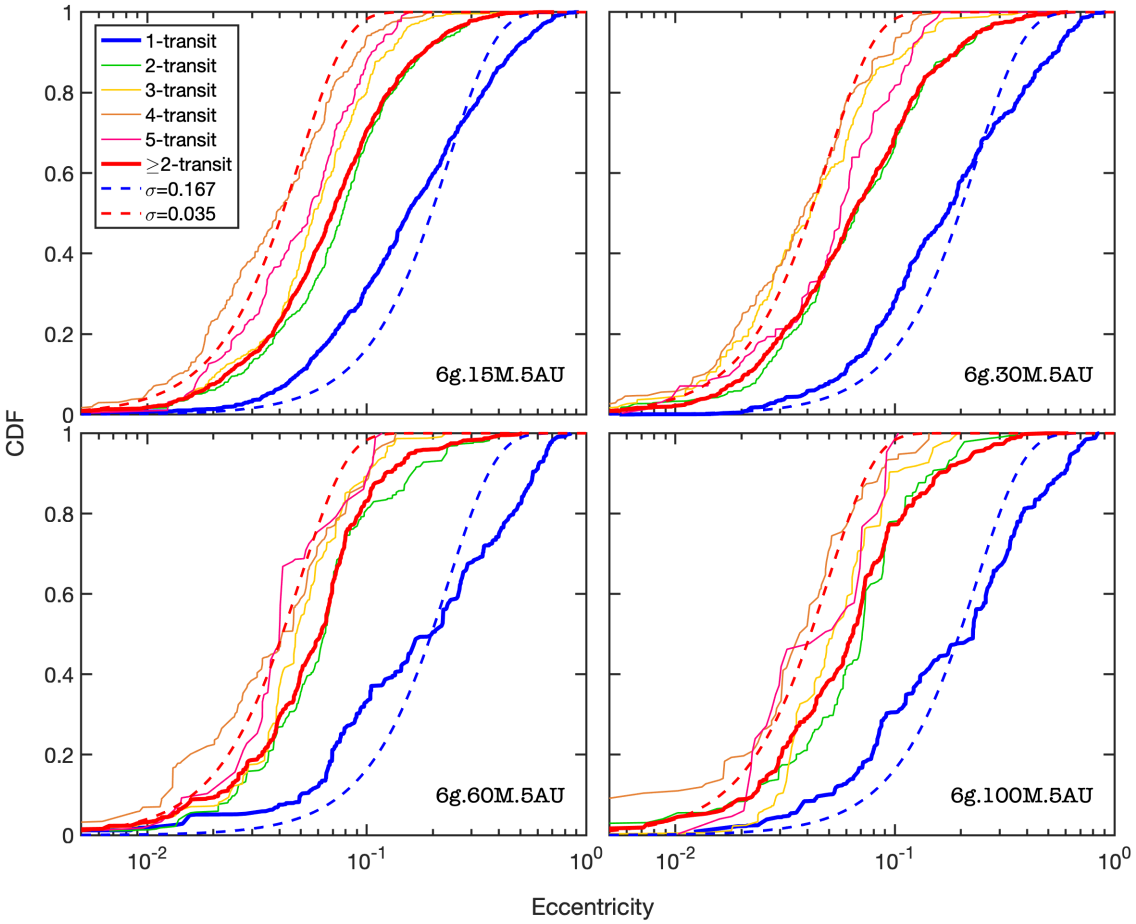

Figure A.14.: CDFs of the eccentricities obtained from the synthetic observation of all outer system templates with $a_{\rm out} = 5$ au. For comparison, the CDFs of eccentricities drawn from Rayleigh distributions with eccentricity parameters $\sigma = 0.167$ and $\sigma = 0.035$ are also plotted.

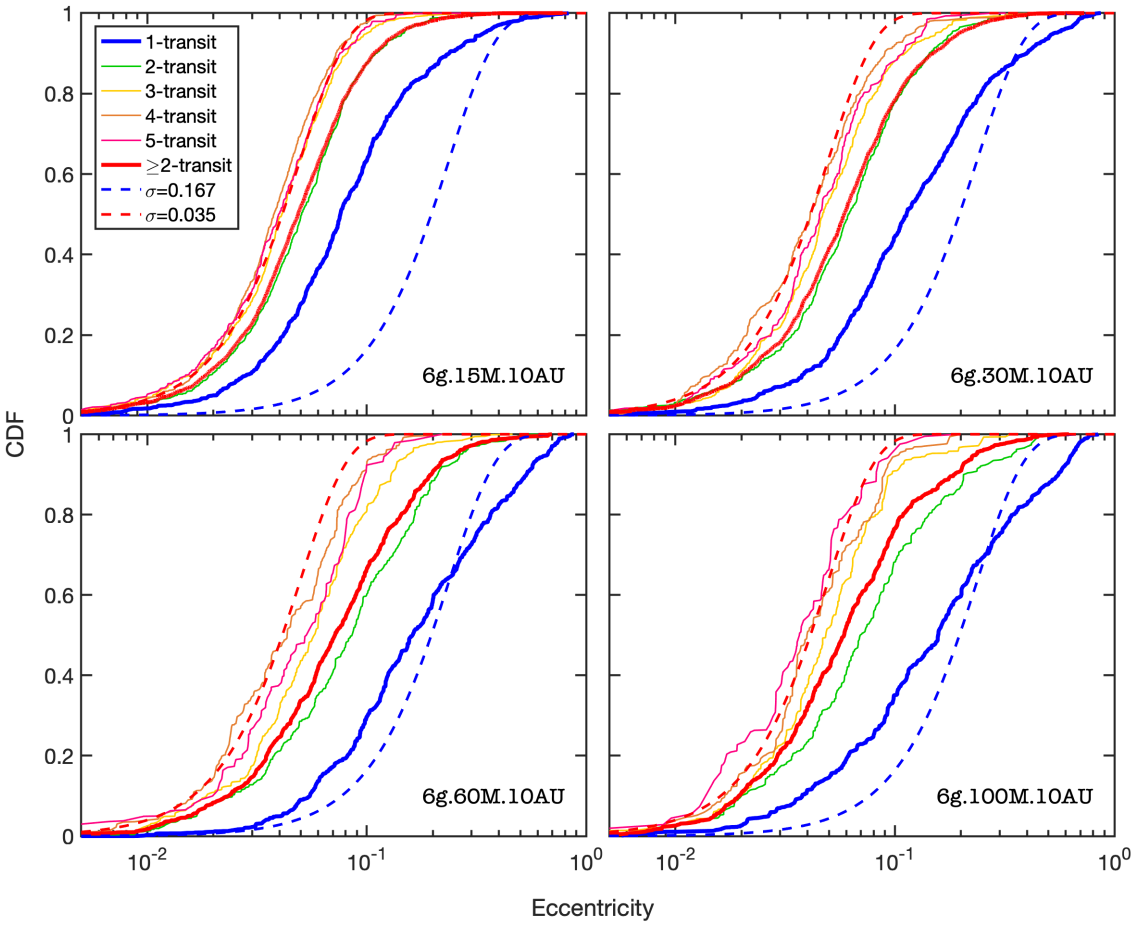

Figure A.15.: CDFs of the eccentricities obtained from the synthetic observation of all outer system templates with $a_{\rm out} = 10$ au. For comparison, the CDFs of eccentricities drawn from Rayleigh distributions with eccentricity parameters $\sigma = 0.167$ and $\sigma = 0.035$ are also plotted.

A.6. Mass-radius relation

Mass-radius relation that we adopted in this study is

$$R_{\rm p} = c_1 \left(\frac{M_{\rm p}}{M_\oplus} \right)^{c_2} R_\oplus, \tag{A.6}$$

where c_1 and c_2 are the mass-radius relation coefficients and are different at different domain. The values of c_1 and c_2 are listed in table A.3. Figure A.16 shows the mass-radius relation, also in comparison with other astronomical objects.

Table A.3.: The value of c_1 and c_2 applied to the mass-radius relation (equation A.6).

Planet mass range [$M_\oplus$]	c_1	c_2
$M_{\rm p} \leqslant 0.01$	1.1748	0.3333
$0.01 < M_{\rm p} \leqslant 2$	1.0014	0.2987
$2 < M_{\rm p} \leqslant 5.8$	1.0340	0.2524
$5.8 < M_{\rm p} \leqslant 91$	0.4372	0.7440
$91 < M_{\rm p} \leqslant 1000$	14.9562	-0.0386

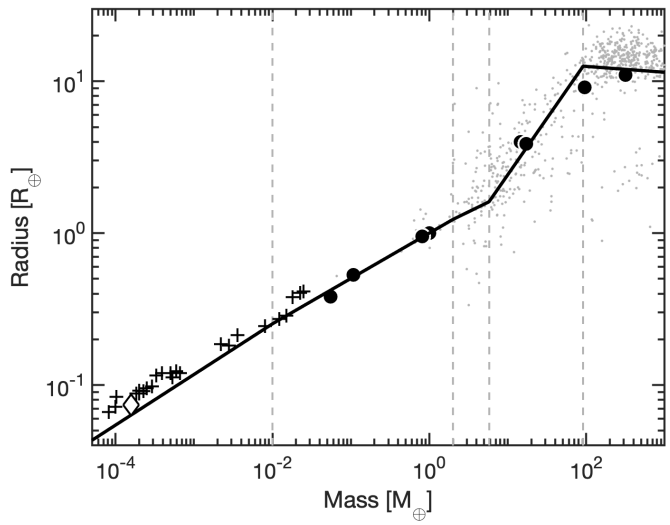

Figure A.16.: Mass-radius relation that adopted in this study, calculated by equation A.6. Comparison to other astronomical objects is provided, including the Solar System planets (black dots), confirmed exoplanets (grey dots), other Solar System objects larger than 400 km (+), and Ceres (♢). Grey vertical dashed lines denote the different planetary mass range.

A.7. System parameters

The system parameters of the five selected systems that we focused in this study, including WASP-47 (Becker et al., 2015), Kepler-730 (Cañas et al., 2019), TOI-1130 (Huang et al., 2020), Kepler-487 (Morton et al., 2016), and Kepler-89 (also known as KOI-94, Weiss et al., 2013).

Table A.4.: Stellar and planet parameters of the five selected systems. All planets in the table are transiting planet and the known masses are given in the form of mass, except WASP-47c which is detected by RV† and the mass is in $M_{\rm p} \sin I$.‡ We list this planet here for the completeness, but the frmation process of this cold giant is not the main consideration in this study. KOI-191.02 and 191.03 are Kepler candidates orbiting their at the Kepler-487 (KOI-191) system*.

		Stellar Parameters		
Host name		$M_\star$ [$M_\odot$]	$R_\star$ [$R_\odot$]	$T_\star$ [K]
WASP-47		1.040	1.137	5552
Kepler-730		1.047	1.411	5620
TOI-1130		0.684	0.687	4250
Kepler-487		0.910	0.880	5444
Kepler-89		1.277	1.520	6182
		Planet Parameters		
Planet name		a [au]	$M_{\rm p}$ [$M_\oplus$]	$R_{\rm p}$ [$R_\oplus$]
WASP-47	e	0.01694	6.83	1.81
-	**b**	**0.05129**	**363.1**	**12.63**
-	d	0.08600	13.1	3.576
-	c†	1.3926	398.2‡	-
Kepler-730	c	0.03997	-	1.57
-	**b**	**0.0694**	-	**12.33**
TOI-1130	b	0.04394	-	3.65
-	**c**	**0.07098**	**309.6**	**16.8**
Kepler-487	191.03*	0.0149	-	1.20
-	191.02*	0.0337	-	2.25
-	**b**	**0.11719**	-	**11.42**
-	c	0.21682	-	2.68
Kepler-89	b	0.05119	10.5	1.71
-	c	0.1013	15.6	4.32
-	**d**	**0.1684**	**106.0**	**11.27**
-	e	0.3046	35.0	6.56

Resources and data usage

This research utilised Queen Mary's Apocrita HPC facility, supported by QMUL Research-IT[1].

This research has made use of the NASA Exoplanet Archive[2], which is operated by the California Institute of Technology, under contract with the National Aeronautics and Space Administration under the Exoplanet Exploration Program.

The data underlying this article will be shared on reasonable request to the corresponding author[3].

[1] http://doi.org/10.5281/zenodo.438045
[2] https://exoplanetarchive.ipac.caltech.edu
[3] s.t.s.poon@qmul.ac.uk

Bibliography

Agnor, C. and E. Asphaug
2004. Accretion Efficiency during Planetary Collisions. *ApJL*, 613(2):L157–L160.

Alfè, D., G. D. Price, and M. J. Gillan
2001. Thermodynamics of hexagonal-close-packed iron under Earth's core conditions. *Phys. Rev. B*, 64(4):045123.

Alibert, Y., C. Mordasini, W. Benz, and C. Winisdoerffer
2005. Models of giant planet formation with migration and disc evolution. *A&A*, 434(1):343–353.

ALMA Partnership, C. L. Brogan, L. M. Pérez, T. R. Hunter, W. R. F. Dent, A. S. Hales, R. E. Hills, S. Corder, E. B. Fomalont, C. Vlahakis, Y. Asaki, D. Barkats, A. Hirota, J. A. Hodge, C. M. V. Impellizzeri, R. Kneissl, E. Liuzzo, R. Lucas, N. Marcelino, S. Matsushita, K. Nakanishi, N. Phillips, A. M. S. Richards, I. Toledo, R. Aladro, D. Broguiere, J. R. Cortes, P. C. Cortes, D. Espada, F. Galarza, D. Garcia-Appadoo, L. Guzman-Ramirez, E. M. Humphreys, T. Jung, S. Kameno, R. A. Laing, S. Leon, G. Marconi, A. Mignano, B. Nikolic, L. A. Nyman, M. Radiszcz, A. Remijan, J. A. Rodón, T. Sawada, S. Takahashi, R. P. J. Tilanus, B. Vila Vilaro, L. C. Watson, T. Wiklind, E. Akiyama, E. Chapillon, I. de Gregorio-Monsalvo, J. Di Francesco, F. Gueth, A. Kawamura, C. F. Lee, Q. Nguyen Luong, J. Mangum, V. Pietu, P. Sanhueza, K. Saigo, S. Takakuwa, C. Ubach, T. van Kempen, A. Wootten, A. Castro-Carrizo, H. Francke, J. Gallardo, J. Garcia, S. Gonzalez, T. Hill, T. Kaminski, Y. Kurono, H. Y. Liu, C. Lopez, F. Morales, K. Plarre, G. Schieven, L. Testi, L. Videla, E. Villard, P. Andreani, J. E. Hibbard, and K. Tatematsu
2015. The 2014 ALMA Long Baseline Campaign: First Results from High Angular Resolution Observations toward the HL Tau Region. *ApJL*, 808(1):L3.

Anderson, K. R., D. Lai, and B. Pu
2020. In situ scattering of warm Jupiters and implications for dynamical histories. *MNRAS*, 491(1):1369–1383.

Anglada-Escudé, G., P. J. Amado, J. Barnes, Z. M. Berdiñas, R. P. Butler, G. A. L. Coleman, I. de La Cueva, S. Dreizler, M. Endl, B. Giesers, S. V. Jeffers, J. S. Jenkins, H. R. A. Jones, M. Kiraga, M. Kürster, M. J. López-González, C. J. Marvin, N. Morales, J. Morin, R. P. Nelson, J. L. Ortiz, A. Ofir, S.-J. Paardekooper, A. Reiners, E. Rodríguez, C. Rodríguez-López, L. F. Sarmiento, J. P. Strachan, Y. Tsapras, M. Tuomi, and M. Zechmeister
2016. A terrestrial planet candidate in a temperate orbit around Proxima Centauri. *Nature*, 536(7617):437–440.

Artymowicz, P.
1993. Disk-Satellite Interaction via Density Waves and the Eccentricity Evolution of Bodies Embedded in Disks. *ApJ*, 419:166.

Asphaug, E.
2010. Similar-sized collisions and the diversity of planets. *Chemie der Erde / Geochemistry*, 70(3):199–219.

Astudillo-Defru, N., R. F. Díaz, X. Bonfils, J. M. Almenara, J. B. Delisle, F. Bouchy, X. Delfosse, T. Forveille, C. Lovis, M. Mayor, F. Murgas, F. Pepe, N. C. Santos, D. Ségransan, S. Udry, and A. Wünsche
2017. The HARPS search for southern extra-solar planets. XLII. A system of Earth-mass planets around the nearby M dwarf YZ Ceti. *A&A*, 605:L11.

Auvergne, M., P. Bodin, L. Boisnard, J.-T. Buey, S. Chaintreuil, G. Epstein, M. Jouret, T. Lam-Trong, P. Levacher, A. Magnan, R. Perez, P. Plasson, J. Plesseria, G. Peter, M. Steller, D. Tiphène, A. Baglin, P. Agogué, T. Appourchaux, D. Barbet, T. Beaufort, R. Bellenger, R. Berlin, P. Bernardi, D. Blouin, P. Boumier, F. Bonneau, R. Briet, B. Butler, R. Cautain, F. Chiavassa, V. Costes, J. Cuvilho, V. Cunha-Parro, F. de Oliveira Fialho, M. Decaudin, J.-M. Defise, S. Djalal, A. Docclo, R. Drummond, O. Dupuis, G. Exil, C. Fauré, A. Gaboriaud, P. Gamet, P. Gavalda, E. Grolleau, L. Gueguen, V. Guivarc'h, P. Guterman, J. Hasiba, G. Huntzinger, H. Hustaix, C. Imbert, G. Jeanville, B. Johlander, L. Jorda, P. Journoud, F. Karioty, L. Kerjean, L. Lafond, V. Lapeyrere, P. Landiech, T. Larqué, P. Laudet, J. Le Merrer, L. Leporati, B. Leruyet, B. Levieuge, A. Llebaria, L. Martin, E. Mazy, J.-M. Mesnager, J.-P. Michel, J.-P. Moalic, W. Monjoin, D. Naudet, S. Neukirchner, K. Nguyen-Kim, M. Ollivier, J.-L. Orcesi, H. Ottacher, A. Oulali, J. Parisot, S. Perruchot, A. Piacentino, L. Pinheiro da Silva, J. Platzer, B. Pontet, A. Pradines, C. Quentin, U. Rohbeck, G. Rolland, F. Rollenhagen, R. Romagnan, N. Russ, R. Samadi, R. Schmidt, N. Schwartz, I. Sebbag, H. Smit, W. Sunter, M. Tello, P. Toulouse, B. Ulmer, O. Vandermarcq, E. Vergnault, R. Wallner, G. Waultier, and P. Zanatta
2009. The CoRoT satellite in flight: description and performance. *A&A*, 506:411–424.

Bailey, E. and K. Batygin
2018. The Hot Jupiter Period-Mass Distribution as a Signature of in situ Formation. *ApJL*, 866(1):L2.

Ballard, S., D. Fabrycky, F. Fressin, D. Charbonneau, J.-M. Desert, G. Torres, G. Marcy, C. J. Burke, H. Isaacson, C. Henze, J. H. Steffen, D. R. Ciardi, S. B. Howell, W. D. Cochran, M. Endl, S. T. Bryson, J. F. Rowe, M. J. Holman, J. J. Lissauer, J. M. Jenkins, M. Still, E. B. Ford, J. L. Christiansen, C. K. Middour, M. R. Haas, J. Li, J. R. Hall, S. McCauliff, N. M. Batalha, D. G. Koch, and W. J. Borucki
2011. The Kepler-19 System: A Transiting 2.2 $R_\oplus$ Planet and a Second Planet Detected via Transit Timing Variations. *ApJ*, 743:200.

Baranne, A., D. Queloz, M. Mayor, G. Adrianzyk, G. Knispel, D. Kohler, D. Lacroix, J. P. Meunier, G. Rimbaud, and A. Vin
1996. ELODIE: A spectrograph for accurate radial velocity measurements. *A&AS*, 119:373–390.

Barclay, T., E. V. Quintana, F. C. Adams, D. R. Ciardi, D. Huber, D. Foreman-Mackey, B. T. Montet, and D. Caldwell
2015. The Five Planets in the Kepler-296 Binary System All Orbit the Primary: A Statistical and Analytical Analysis. *ApJ*, 809:7.

Batalha, N. M., J. F. Rowe, S. T. Bryson, T. Barclay, C. J. Burke, D. A. Caldwell, J. L. Christiansen, F. Mullally, S. E. Thompson, T. M. Brown, A. K. Dupree, D. C. Fabrycky, E. B. Ford, J. J. Fortney, R. L. Gilliland, H. Isaacson, D. W. Latham, G. W. Marcy, S. N. Quinn, D. Ragozzine, A. Shporer, W. J. Borucki, D. R. Ciardi, I. Gautier, Thomas N., M. R. Haas, J. M. Jenkins, D. G. Koch, J. J. Lissauer, W. Rapin, G. S. Basri, A. P. Boss, L. A. Buchhave, J. A. Carter, D. Charbonneau, J. Christensen-Dalsgaard, B. D. Clarke, W. D. Cochran, B.-O. Demory, J.-M. Desert, E. Devore, L. R. Doyle, G. A. Esquerdo, M. Everett, F. Fressin, J. C. Geary, F. R. Girouard, A. Gould, J. R. Hall, M. J. Holman, A. W. Howard, S. B. Howell, K. A. Ibrahim, K. Kinemuchi, H. Kjeldsen, T. C. Klaus, J. Li, P. W. Lucas, S. Meibom, R. L. Morris, A. Prša, E. Quintana, D. T. Sanderfer, D. Sasselov, S. E. Seader, J. C. Smith, J. H. Steffen, M. Still, M. C. Stumpe, J. C. Tarter, P. Tenenbaum, G. Torres, J. D. Twicken, K. Uddin, J. Van Cleve, L. Walkowicz, and W. F. Welsh
2013. Planetary Candidates Observed by Kepler. III. Analysis of the First 16 Months of Data. *The Astrophysical Journal Supplement Series*, 204:24.

Batygin, K., P. H. Bodenheimer, and G. P. Laughlin
2016. In Situ Formation and Dynamical Evolution of Hot Jupiter Systems. *ApJ*, 829(2):114.

Beaulieu, J.-P., D. P. Bennett, P. Fouqué, A. Williams, M. Dominik, U. G. Jørgensen, D. Kubas, A. Cassan, C. Coutures, J. Greenhill, K. Hill, J. Menzies, P. D. Sackett, M. Albrow, S. Brillant, J. A. R. Caldwell, J. J. Calitz, K. H. Cook, E. Corrales, M. Desort, S. Dieters, D. Dominis, J. Donatowicz, M. Hoffman, S. Kane, J.-B. Marquette, R. Martin, P. Meintjes, K. Pollard, K. Sahu, C. Vinter, J. Wambsganss, K. Woller, K. Horne, I. Steele, D. M. Bramich, M. Burgdorf, C. Snodgrass, M. Bode, A. Udalski, M. K. Szymański, M. Kubiak, T. Więckowski, G. Pietrzyński, I. Soszyński, O. Szewczyk, Ł. Wyrzykowski, B. Paczyński, F. Abe, I. A. Bond, T. R. Britton, A. C. Gilmore, J. B. Hearnshaw, Y. Itow, K. Kamiya, P. M. Kilmartin, A. V. Korpela, K. Masuda, Y. Matsubara, M. Motomura, Y. Muraki, S. Nakamura, C. Okada, K. Ohnishi, N. J. Rattenbury, T. Sako, S. Sato, M. Sasaki, T. Sekiguchi, D. J. Sullivan, P. J. Tristram, P. C. M. Yock, and T. Yoshioka
2006. Discovery of a cool planet of 5.5 Earth masses through gravitational microlensing. *Nature*, 439:437–440.

Becker, J. C. and F. C. Adams
2017. Effects of unseen additional planetary perturbers on compact extrasolar planetary systems. *MNRAS*, 468(1):549–563.

Becker, J. C., A. Vanderburg, F. C. Adams, S. A. Rappaport, and H. M. Schwengeler
2015. WASP-47: A Hot Jupiter System with Two Additional Planets Discovered by K2. *ApJL*, 812(2):L18.

Beckers, J. M.
1993. Adaptive optics for astronomy - Principles, performance, and applications. *ARAA*, 31:13–62.

Biersteker, J. B. and H. E. Schlichting
2019. Atmospheric mass-loss due to giant impacts: the importance of the thermal component for hydrogen-helium envelopes. *MNRAS*, 485(3):4454–4463.

Bitsch, B., A. Izidoro, A. Johansen, S. N. Raymond, A. Morbidelli, M. Lambrechts, and S. A. Jacobson
2019. Formation of planetary systems by pebble accretion and migration: growth of gas giants. *A&A*, 623:A88.

Bitsch, B., M. Lambrechts, and A. Johansen
2015. The growth of planets by pebble accretion in evolving protoplanetary discs. *A&A*, 582:A112.

Bodenheimer, P. and J. J. Lissauer
2014. Accretion and Evolution of ~2.5 M $_\oplus$ Planets with Voluminous H/He Envelopes. *ApJ*, 791(2):103.

Bodenheimer, P. and J. B. Pollack
1986. Calculations of the accretion and evolution of giant planets: The effects of solid cores. *Icarus*, 67(3):391–408.

Boley, A. C., A. P. Granados Contreras, and B. Gladman
2016. The In Situ Formation of Giant Planets at Short Orbital Periods. *ApJL*, 817(2):L17.

Bond, I. A., N. J. Rattenbury, J. Skuljan, F. Abe, R. J. Dodd, J. B. Hearnshaw, M. Honda, J. Jugaku, P. M. Kilmartin, A. Marles, K. Masuda, Y. Matsubara, Y. Muraki, T. Nakamura, G. Nankivell, S. Noda, C. Noguchi, K. Ohnishi, M. Reid, T. Saito, H. Sato, M. Sekiguchi, D. J. Sullivan, T. Sumi, M. Takeuti, Y. Watase, S. Wilkinson, R. Yamada, T. Yanagisawa, and P. C. M. Yock
2002. Study by MOA of extrasolar planets in gravitational microlensing events of high magnification. *MNRAS*, 333:71–83.

Borucki, W., D. Koch, A. Boss, E. Dunham, A. Dupree, J. Geary, R. Gilliland, S. Howell, J. Jenkins, Y. Kondo, D. Latham, J. Lissauer, and H. Reitsema
2004. The Kepler mission: a technical overview. In *Stellar Structure and Habitable*

Planet Finding, F. Favata, S. Aigrain, and A. Wilson, eds., volume 538 of *ESA Special Publication*, Pp. 177–182.

Borucki, W. J., D. Koch, G. Basri, N. Batalha, T. Brown, D. Caldwell, J. Caldwell, J. Christensen-Dalsgaard, W. D. Cochran, E. DeVore, E. W. Dunham, A. K. Dupree, T. N. Gautier, J. C. Geary, R. Gilliland, A. Gould, S. B. Howell, J. M. Jenkins, Y. Kondo, D. W. Latham, G. W. Marcy, S. Meibom, H. Kjeldsen, J. J. Lissauer, D. G. Monet, D. Morrison, D. Sasselov, J. Tarter, A. Boss, D. Brownlee, T. Owen, D. Buzasi, D. Charbonneau, L. Doyle, J. Fortney, E. B. Ford, M. J. Holman, S. Seager, J. H. Steffen, W. F. Welsh, J. Rowe, H. Anderson, L. Buchhave, D. Ciardi, L. Walkowicz, W. Sherry, E. Horch, H. Isaacson, M. E. Everett, D. Fischer, G. Torres, J. A. Johnson, M. Endl, P. MacQueen, S. T. Bryson, J. Dotson, M. Haas, J. Kolodziejczak, J. Van Cleve, H. Chandrasekaran, J. D. Twicken, E. V. Quintana, B. D. Clarke, C. Allen, J. Li, H. Wu, P. Tenenbaum, E. Verner, F. Bruhweiler, J. Barnes, and A. Prsa
2010. Kepler Planet-Detection Mission: Introduction and First Results. *Science*, 327:977.

Borucki, W. J., D. G. Koch, G. Basri, N. Batalha, T. M. Brown, S. T. Bryson, D. Caldwell, J. Christensen-Dalsgaard, W. D. Cochran, E. DeVore, E. W. Dunham, I. Gautier, Thomas N., J. C. Geary, R. Gilliland, A. Gould, S. B. Howell, J. M. Jenkins, D. W. Latham, J. J. Lissauer, G. W. Marcy, J. Rowe, D. Sasselov, A. Boss, D. Charbonneau, D. Ciardi, L. Doyle, A. K. Dupree, E. B. Ford, J. Fortney, M. J. Holman, S. Seager, J. H. Steffen, J. Tarter, W. F. Welsh, C. Allen, L. A. Buchhave, J. L. Christiansen, B. D. Clarke, S. Das, J.-M. Désert, M. Endl, D. Fabrycky, F. Fressin, M. Haas, E. Horch, A. Howard, H. Isaacson, H. Kjeldsen, J. Kolodziejczak, C. Kulesa, J. Li, P. W. Lucas, P. Machalek, D. McCarthy, P. MacQueen, S. Meibom, T. Miquel, A. Prsa, S. N. Quinn, E. V. Quintana, D. Ragozzine, W. Sherry, A. Shporer, P. Tenenbaum, G. Torres, J. D. Twicken, J. Van Cleve, L. Walkowicz, F. C. Witteborn, and M. Still
2011. Characteristics of Planetary Candidates Observed by Kepler. II. Analysis of the First Four Months of Data. *ApJ*, 736:19.

Borucki, W. J., D. G. Koch, N. Batalha, S. T. Bryson, J. Rowe, F. Fressin, G. Torres, D. A. Caldwell, J. Christensen-Dalsgaard, W. D. Cochran, E. DeVore, T. N. Gautier, J. C. Geary, R. Gilliland, A. Gould, S. B. Howell, J. M. Jenkins, D. W. Latham, J. J. Lissauer, G. W. Marcy, D. Sasselov, A. Boss, D. Charbonneau, D. Ciardi, L. Kaltenegger, L. Doyle, A. K. Dupree, E. B. Ford, J. Fortney, M. J. Holman, J. H. Steffen, F. Mullally, M. Still, J. Tarter, S. Ballard, L. A. Buchhave, J. Carter, J. L. Christiansen, B.-O. Demory, J.-M. Désert, C. Dressing, M. Endl, D. Fabrycky, D. Fischer, M. R. Haas, C. Henze, E. Horch, A. W. Howard, H. Isaacson, H. Kjeldsen, J. A. Johnson, T. Klaus, J. Kolodziejczak, T. Barclay, J. Li, S. Meibom, A. Prsa, S. N. Quinn, E. V. Quintana, P. Robertson, W. Sherry, A. Shporer, P. Tenenbaum, S. E. Thompson, J. D. Twicken, J. Van Cleve, W. F. Welsh, S. Basu, W. Chaplin, A. Miglio, S. D. Kawaler, T. Arentoft, D. Stello, T. S. Metcalfe, G. A. Verner, C. Karoff, M. Lundkvist, M. N. Lund, R. Handberg,

Y. Elsworth, S. Hekker, D. Huber, T. R. Bedding, and W. Rapin
2012. Kepler-22b: A 2.4 Earth-radius Planet in the Habitable Zone of a Sun-like Star. *ApJ*, 745:120.

Boss, A.
2009. *The crowded universe : the search for living planets*. New York : Basic Books.

Boss, A. P.
1997. Giant planet formation by gravitational instability. *Science*, 276:1836–1839.

Brown, T. M., D. W. Latham, M. E. Everett, and G. A. Esquerdo
2011. Kepler Input Catalog: Photometric Calibration and Stellar Classification. *AJ*, 142:112.

Burke, C. J., S. T. Bryson, F. Mullally, J. F. Rowe, J. L. Christiansen, S. E. Thompson, J. L. Coughlin, M. R. Haas, N. M. Batalha, D. A. Caldwell, J. M. Jenkins, M. Still, T. Barclay, W. J. Borucki, W. J. Chaplin, D. R. Ciardi, B. D. Clarke, W. D. Cochran, B.-O. Demory, G. A. Esquerdo, I. Gautier, Thomas N., R. L. Gilliland, F. R. Girouard, M. Havel, C. E. Henze, S. B. Howell, D. Huber, D. W. Latham, J. Li, R. C. Morehead, T. D. Morton, J. Pepper, E. Quintana, D. Ragozzine, S. E. Seader, Y. Shah, A. Shporer, P. Tenenbaum, J. D. Twicken, and A. Wolfgang
2014. Planetary Candidates Observed by Kepler IV: Planet Sample from Q1-Q8 (22 Months). *The Astrophysical Journal Supplement Series*, 210:19.

Burns, J. A., P. L. Lamy, and S. Soter
1979. Radiation forces on small particles in the solar system. *Icarus*, 40:1–48.

Butler, R. P., J. T. Wright, G. W. Marcy, D. A. Fischer, S. S. Vogt, C. G. Tinney, H. R. A. Jones, B. D. Carter, J. A. Johnson, C. McCarthy, and A. J. Penny
2006. Catalog of Nearby Exoplanets. *ApJ*, 646(1):505–522.

Cañas, C. I., S. Wang, S. Mahadevan, C. F. Bender, N. De Lee, S. W. Fleming, D. A. García-Hernández, F. R. Hearty, S. R. Majewski, A. Roman-Lopes, D. P. Schneider, and K. G. Stassun
2019. Kepler-730: A Hot Jupiter System with a Close-in, Transiting, Earth-sized Planet. *ApJL*, 870(2):L17.

Cabrera, J., S. Csizmadia, H. Lehmann, R. Dvorak, D. Gandolfi, H. Rauer, A. Erikson, C. Dreyer, P. Eigmüller, and A. Hatzes
2014. The Planetary System to KIC 11442793: A Compact Analogue to the Solar System. *ApJ*, 781:18.

Cameron, A. G. W.
1978. Physics of the Primitive Solar Accretion Disk. *Moon and Planets*, 18(1):5–40.

Campante, T. L., T. Barclay, J. J. Swift, D. Huber, V. Z. Adibekyan, W. Cochran, C. J. Burke, H. Isaacson, E. V. Quintana, G. R. Davies, V. Silva Aguirre, D. Ragozzine, R. Riddle, C. Baranec, S. Basu, W. J. Chaplin, J. Christensen-Dalsgaard, T. S. Metcalfe, T. R. Bedding, R. Handberg, D. Stello, J. M. Brewer, S. Hekker, C. Karoff, R. Kolbl, N. M. Law, M. Lundkvist, A. Miglio, J. F. Rowe, N. C. Santos, C. Van Laerhoven, T. Arentoft, Y. P. Elsworth, D. A. Fischer, S. D. Kawaler, H. Kjeldsen, M. N. Lund, G. W. Marcy, S. G. Sousa, A. Sozzetti, and T. R. White
2015. An Ancient Extrasolar System with Five Sub-Earth-size Planets. *ApJ*, 799(2):170.

Campbell, B., G. A. H. Walker, and S. Yang
1988. A search for substellar companions to solar-type stars. *ApJ*, 331:902–921.

Carrera, D., E. B. Ford, and A. Izidoro
2019. Formation of short-period planets by disc migration. *MNRAS*, 486(3):3874–3885.

Carter, P. J., S. J. Lock, and S. T. Stewart
2020. The Energy Budgets of Giant Impacts. *Journal of Geophysical Research (Planets)*, 125(1):e06042.

Casertano, S., M. G. Lattanzi, A. Sozzetti, A. Spagna, S. Jancart, R. Morbidelli, R. Pannunzio, D. Pourbaix, and D. Queloz
2008. Double-blind test program for astrometric planet detection with Gaia. *A&A*, 482:699–729.

Chambers, J. E.
1999. A hybrid symplectic integrator that permits close encounters between massive bodies. *MNRAS*, 304:793–799.

Charpinet, S., G. Fontaine, P. Brassard, E. M. Green, V. Van Grootel, S. K. Randall, R. Silvotti, A. S. Baran, R. H. Østensen, S. D. Kawaler, and J. H. Telting
2011. A compact system of small planets around a former red-giant star. *Nature*, 480:496–499.

Chatterjee, S., E. B. Ford, S. Matsumura, and F. A. Rasio
2008. Dynamical Outcomes of Planet-Planet Scattering. *ApJ*, 686(1):580–602.

Chiang, E. and G. Laughlin
2013. The minimum-mass extrasolar nebula: in situ formation of close-in super-Earths. *MNRAS*, 431(4):3444–3455.

Clarke, C. J., A. Gendrin, and M. Sotomayor
2001. The dispersal of circumstellar discs: the role of the ultraviolet switch. *MNRAS*, 328(2):485–491.

Cochran, W. D., A. P. Hatzes, and T. J. Hancock
1991. Constraints on the companion object to HD 114762. *ApJL*, 380:L35–L38.

Coleman, G. A. L. and R. P. Nelson
2014. On the formation of planetary systems via oligarchic growth in thermally evolving viscous discs. *MNRAS*, 445(1):479–499.

Coleman, G. A. L. and R. P. Nelson
2016a. Giant planet formation in radially structured protoplanetary discs. *MNRAS*, 460:2779–2795.

Coleman, G. A. L. and R. P. Nelson
2016b. On the formation of compact planetary systems via concurrent core accretion and migration. *MNRAS*, 457:2480–2500.

Coleman, G. A. L., J. C. B. Papaloizou, and R. P. Nelson
2017. In situ accretion of gaseous envelopes on to planetary cores embedded in evolving protoplanetary discs. *MNRAS*, 470(3):3206–3219.

Cossou, C., S. N. Raymond, F. Hersant, and A. Pierens
2014. Hot super-Earths and giant planet cores from different migration histories. *A&A*, 569:A56.

Coughlin, J. L., F. Mullally, S. E. Thompson, J. F. Rowe, C. J. Burke, D. W. Latham, N. M. Batalha, A. Ofir, B. L. Quarles, C. E. Henze, A. Wolfgang, D. A. Caldwell, S. T. Bryson, A. Shporer, J. Catanzarite, R. Akeson, T. Barclay, W. J. Borucki, T. S. Boyajian, J. R. Campbell, J. L. Christiansen, F. R. Girouard, M. R. Haas, S. B. Howell, D. Huber, J. M. Jenkins, J. Li, A. Patil-Sabale, E. V. Quintana, S. Ramirez, S. Seader, J. C. Smith, P. Tenenbaum, J. D. Twicken, and K. A. Zamudio
2016. Planetary Candidates Observed by Kepler. VII. The First Fully Uniform Catalog Based on the Entire 48-month Data Set (Q1-Q17 DR24). *The Astrophysical Journal Supplement Series*, 224:12.

Cresswell, P. and R. P. Nelson
2006. On the evolution of multiple protoplanets embedded in a protostellar disc. *A&A*, 450(2):833–853.

Cresswell, P. and R. P. Nelson
2008. Three-dimensional simulations of multiple protoplanets embedded in a protostellar disc. *A&A*, 482:677–690.

Crida, A., A. Morbidelli, and F. Masset
2006. On the width and shape of gaps in protoplanetary disks. *Icarus*, 181(2):587–604.

da Silva, R., S. Udry, F. Bouchy, M. Mayor, C. Moutou, F. Pont, D. Queloz, N. C. Santos, D. Ségransan, and S. Zucker
2006. Elodie metallicity-biased search for transiting Hot Jupiters. I. Two Hot Jupiters orbiting the slightly evolved stars <ASTROBJ>HD 118203</ASTROBJ> and <ASTROBJ>HD 149143</ASTROBJ>. *A&A*, 446(2):717–722.

Dawson, R. I. and J. A. Johnson
2018. Origins of Hot Jupiters. *ARAA*, 56:175–221.

de Laplace, P. S.
1796. *Exposition du système du monde.*

Dekker, H., S. D'Odorico, A. Kaufer, B. Delabre, and H. Kotzlowski
2000. Design, construction, and performance of UVES, the echelle spectrograph for the UT2 Kueyen Telescope at the ESO Paranal Observatory. In *Optical and IR Telescope Instrumentation and Detectors*, M. Iye and A. F. Moorwood, eds., volume 4008 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Pp. 534–545.

Dermott, S. F. and C. D. Murray
1981a. The dynamics of tadpole and horseshoe orbits I. Theory. *Icarus*, 48:1–11.

Dermott, S. F. and C. D. Murray
1981b. The dynamics of tadpole and horseshoe orbits II. The coorbital satellites of Saturn. *Icarus*, 48:12–22.

Dominik, M.
2010. Studying planet populations with Einstein's blip. *Philosophical Transactions of the Royal Society of London Series A*, 368:3535–3550.

Dressing, C. D. and D. Charbonneau
2013. The Occurrence Rate of Small Planets around Small Stars. *ApJ*, 767(1):95.

Dressing, C. D. and D. Charbonneau
2015. The Occurrence of Potentially Habitable Planets Orbiting M Dwarfs Estimated from the Full Kepler Dataset and an Empirical Measurement of the Detection Sensitivity. *ApJ*, 807(1):45.

Dullemond, C. P., D. Hollenbach, I. Kamp, and P. D'Alessio
2007. Models of the Structure and Evolution of Protoplanetary Disks. In *Protostars and Planets V*, B. Reipurth, D. Jewitt, and K. Keil, eds., P. 555.

Duncan, M. J., H. F. Levison, and M. H. Lee
1998. A Multiple Time Step Symplectic Algorithm for Integrating Close Encounters. *AJ*, 116:2067–2077.

Dunham, E. W., W. J. Borucki, D. G. Koch, N. M. Batalha, L. A. Buchhave, T. M. Brown, D. A. Caldwell, W. D. Cochran, M. Endl, D. Fischer, G. Fűrész, I. Gautier, Thomas N., J. C. Geary, R. L. Gilliland , A. Gould, S. B. Howell, J. M. Jenkins, H. Kjeldsen, D. W. Latham, J. J. Lissauer, G. W. Marcy, S. Meibom, D. G. Monet, J. F. Rowe, and D. D. Sasselov
2010. Kepler-6b: A Transiting Hot Jupiter Orbiting a Metal-rich Star. *ApJL*, 713(2):L136–L139.

Einstein, A.
1916. Die Grundlage der allgemeinen Relativitätstheorie. *Annalen der Physik*, 354:769–822.

Flagg, L., C. M. Johns-Krull, L. Nofi, J. Llama, L. Prato, K. Sullivan, D. T. Jaffe, and G. Mace
2019. CO Detected in CI Tau b: Hot Start Implied by Planet Mass and M_K. *ApJL*, 878(2):L37.

Fogg, M. J. and R. P. Nelson
2005. Oligarchic and giant impact growth of terrestrial planets in the presence of gas giant planet migration. *A&A*, 441(2):791–806.

Fogg, M. J. and R. P. Nelson
2007. On the formation of terrestrial planets in hot-Jupiter systems. *A&A*, 461(3):1195–1208.

Ford, E. B., S. N. Quinn, and D. Veras
2008. Characterizing the Orbital Eccentricities of Transiting Extrasolar Planets with Photometric Observations. *ApJ*, 678(2):1407–1418.

Ford, E. B. and F. A. Rasio
2008. Origins of Eccentric Extrasolar Planets: Testing the Planet-Planet Scattering Model. *ApJ*, 686(1):621–636.

Fressin, F., G. Torres, D. Charbonneau, S. T. Bryson, J. Christiansen, C. D. Dressing, J. M. Jenkins, L. M. Walkowicz, and N. M. Batalha
2013. The False Positive Rate of Kepler and the Occurrence of Planets. *ApJ*, 766(2):81.

Freudenthal, J., C. von Essen, A. Ofir, S. Dreizler, E. Agol, S. Wedemeyer, B. M. Morris, A. C. Becker, H. J. Deeg, S. Hoyer, M. Mallonn, K. Poppenhaeger, E. Herrero, I. Ribas, P. Boumis, and A. Liakos
2019. Kepler Object of Interest Network. III. Kepler-82f: a new non-transiting 21 $M_\oplus$ planet from photodynamical modelling. *A&A*, 628:A108.

Fulton, B. J. and E. A. Petigura
2018. The California-Kepler Survey. VII. Precise Planet Radii Leveraging Gaia DR2 Reveal the Stellar Mass Dependence of the Planet Radius Gap. *AJ*, 156(6):264.

Fulton, B. J., E. A. Petigura, A. W. Howard, H. Isaacson, G. W. Marcy, P. A. Cargile, L. Hebb, L. M. Weiss, J. A. Johnson, T. D. Morton, E. Sinukoff, I. J. M. Crossfield, and L. A. Hirsch
2017. The California-Kepler Survey. III. A Gap in the Radius Distribution of Small Planets. *AJ*, 154(3):109.

Funk, B., R. Dvorak, and R. Schwarz
2013. Exchange orbits: an interesting case of co-orbital motion. *Celestial Mechanics and Dynamical Astronomy*, 117:41–58.

Funk, B., R. Schwarz, R. Dvorak, and M. Roth
2011. Exchange orbits: a possible application to extrasolar planetary systems? *MNRAS*, 410:455–460.

Gammie, C. F.
2001. Nonlinear Outcome of Gravitational Instability in Cooling, Gaseous Disks. *ApJ*, 553(1):174–183.

Gautier, Thomas N., I., N. M. Batalha, W. J. Borucki, W. D. Cochran, E. W. Dunham, S. B. Howell, D. G. Koch, D. W. Latham, G. W. Marcy, L. A. Buchhave, D. R. Ciardi, M. Endl, G. Furesz, H. Isaacson, P. MacQueen, G. Mandushev, and L. Walkowicz
2010. The Kepler Follow-up Observation Program. *arXiv e-prints*.

Genda, H. and Y. Abe
2003. Survival of a proto-atmosphere through the stage of giant impacts: the mechanical aspects. *Icarus*, 164(1):149–162.

Genda, H., E. Kokubo, and S. Ida
2012. Merging Criteria for Giant Impacts of Protoplanets. *ApJ*, 744(2):137.

Gilliland, R. L., G. W. Marcy, J. F. Rowe, L. Rogers, G. Torres, F. Fressin, E. D. Lopez, L. A. Buchhave, J. Christensen-Dalsgaard, J.-M. Désert, C. E. Henze, H. Isaacson, J. M. Jenkins, J. J. Lissauer, W. J. Chaplin, S. Basu, T. S. Metcalfe, Y. Elsworth, R. Handberg, S. Hekker, D. Huber, C. Karoff, H. Kjeldsen, M. N. Lund, M. Lundkvist, A. Miglio, D. Charbonneau, E. B. Ford, J. J. Fortney, M. R. Haas, A. W. Howard, S. B. Howell, D. Ragozzine, and S. E. Thompson
2013. Kepler-68: Three Planets, One with a Density between that of Earth and Ice Giants. *ApJ*, 766(1):40.

Goldreich, P. and S. Soter
1966. Q in the Solar System. *Icarus*, 5(1):375–389.

Goldreich, P. and S. Tremaine
1980. Disk-satellite interactions. *ApJ*, 241:425–441.

Guess, A. W.
1962. Poynting-Robertson effect for a spherical source of radiation. *ApJ*, 135:855–866.

Hadden, S. and Y. Lithwick
2014. Densities and Eccentricities of 139 Kepler Planets from Transit Time Variations. *ApJ*, 787:80.

Haisch, Karl E., J., E. A. Lada, and C. J. Lada
2001. Disk Frequencies and Lifetimes in Young Clusters. *ApJL*, 553(2):L153–L156.

Hansen, B. M. S.
2017. Perturbation of Compact Planetary Systems by Distant Giant Planets. *MNRAS*, 467(2):1531–1560.

Hansen, B. M. S. and N. Murray
2012. Migration Then Assembly: Formation of Neptune-mass Planets inside 1 AU. *ApJ*, 751(2):158.

Hansen, B. M. S. and N. Murray
2013. Testing in Situ Assembly with the Kepler Planet Candidate Sample. *ApJ*, 775:53.

Hasegawa, Y., T. Y. M. Yu, and B. M. S. Hansen
2019. Close-in giant-planet formation via in-situ gas accretion and their natal disk properties. *A&A*, 629:L1.

Hatzes, A. P., W. D. Cochran, M. Endl, B. McArthur, D. B. Paulson, G. A. H. Walker, B. Campbell, and S. Yang
2003. A Planetary Companion to γ Cephei A. *ApJ*, 599(2):1383–1394.

Hayashi, C.
1981. Structure of the Solar Nebula, Growth and Decay of Magnetic Fields and Effects of Magnetic and Turbulent Viscosities on the Nebula. *Progress of Theoretical Physics Supplement*, 70:35–53.

Hébrard, G., R. F. Díaz, A. C. M. Correia, A. Collier Cameron, J. Laskar, D. Pollacco, J. M. Almenara, D. R. Anderson, S. C. C. Barros, I. Boisse, A. S. Bonomo, F. Bouchy, G. Boué, P. Boumis, D. J. A. Brown, S. Dalal, M. Deleuil, O. D. S. Demangeon, A. P. Doyle, C. A. Haswell, C. Hellier, H. Osborn, F. Kiefer, U. C. Kolb, K. Lam, A. Lecavelier des Étangs, T. Lopez, M. Martin-Lagarde, P. Maxted, J. McCormac, L. D. Nielsen, E. Pallé, J. Prieto-Arranz, D. Queloz, A. Santerne, B. Smalley, O. Turner, S. Udry, D. Verilhac, R. West, P. J. Wheatley, and P. A. Wilson
2020. Discovery and characterization of the exoplanets WASP-148b and c. A transiting system with two interacting giant planets. *A&A*, 640:A32.

Hellary, P. and R. P. Nelson
2012. Global models of planetary system formation in radiatively-inefficient protoplanetary discs. *MNRAS*, 419(4):2737–2757.

Hollenbach, D., D. Johnstone, S. Lizano, and F. Shu
1994. Photoevaporation of Disks around Massive Stars and Application to Ultracompact H II Regions. *ApJ*, 428:654.

Holsapple, K. A. and R. M. Schmidt
1987. Point source solutions and coupling parameters in cratering mechanics. *J. Geophys. Res.*, 92:6350–6376.

Housen, K. R. and K. A. Holsapple
1990. On the fragmentation of asteroids and planetary satellites. *Icarus*, 84(1):226–253.

Howard, A. W., J. A. Johnson, G. W. Marcy, D. A. Fischer, J. T. Wright, D. Bernat, G. W. Henry, K. M. G. Peek, H. Isaacson, K. Apps, M. Endl, W. D. Cochran, J. A. Valenti, J. Anderson, and N. E. Piskunov
2010. The California Planet Survey. I. Four New Giant Exoplanets. *ApJ*, 721(2):1467–1481.

Huang, C. X., C. Petrovich, and E. Deibert
2017. Dynamically Hot Super-Earths from Outer Giant Planet Scattering. *AJ*, 153(5):210.

Huang, C. X., S. N. Quinn, A. Vanderburg, J. Becker, J. E. Rodriguez, F. J. Pozuelos, D. Gandolfi, G. Zhou, A. W. Mann, K. A. Collins, I. Crossfield, K. Barkaoui, K. I. Collins, M. Fridlund, M. Gillon, E. J. Gonzales, M. N. Günther, T. J. Henry, S. B. Howell, H.-S. James, W.-C. Jao, E. Jehin, E. L. N. Jensen, S. R. Kane, J. J. Lissauer, E. Matthews, R. A. Matson, L. A. Paredes, J. E. Schlieder, K. G. Stassun, A. Shporer, L. Sha, T.-G. Tan, I. Georgieva, S. Mathur, E. Palle, C. M. Persson, V. Van Eylen, G. R. Ricker, R. K. Vanderspek, D. W. Latham, J. N. Winn, S. Seager, J. M. Jenkins, C. J. Burke, R. F. Goeke, S. Rinehart, M. E. Rose, E. B. Ting, G. Torres, and I. Wong
2020. TESS Spots a Hot Jupiter with an Inner Transiting Neptune. *ApJL*, 892(1):L7.

Hwang, K. H., A. Udalski, Y. Shvartzvald, Y. H. Ryu, M. D. Albrow, S. J. Chung, A. Gould, C. Han, Y. K. Jung, I. G. Shin, J. C. Yee, W. Zhu, S. M. Cha, D. J. Kim, H. W. Kim, S. L. Kim, C. U. Lee, D. J. Lee, Y. Lee, B. G. Park, R. W. Pogge, KMTNet Collaboration, J. Skowron, P. Mróz, R. Poleski, S. Kozłowski, I. Soszyński, P. Pietrukowicz, M. K. Szymański, K. Ulaczyk, M. Pawlak, OGLE Collaboration, G. Bryden, C. Beichman, S. Calchi Novati, B. S. Gaudi, C. B. Henderson, S. Jacklin, M. T. Penny, and UKIRT Microlensing Team
2018. OGLE-2017-BLG-0173Lb: Low-mass-ratio Planet in a "Hollywood" Microlensing Event. *AJ*, 155(1):20.

Ida, S., T. Guillot, and A. Morbidelli
2016. The radial dependence of pebble accretion rates: A source of diversity in planetary systems. I. Analytical formulation. *A&A*, 591:A72.

Inamdar, N. K. and H. E. Schlichting
2016. Stealing the Gas: Giant Impacts and the Large Diversity in Exoplanet Densities. *ApJL*, 817(2):L13.

Izidoro, A., B. Bitsch, S. N. Raymond, A. Johansen, A. Morbidelli, M. Lambrechts, and S. A. Jacobson
2019. Formation of planetary systems by pebble accretion and migration: Hot super-Earth systems from breaking compact resonant chains. *arXiv e-prints*, P. arXiv:1902.08772.

Izidoro, A., M. Ogihara, S. N. Raymond, A. Morbidelli, A. Pierens, B. Bitsch, C. Cossou, and F. Hersant
2017. Breaking the chains: hot super-Earth systems from migration and disruption of compact resonant chains. *MNRAS*, 470:1750–1770.

Jackson, B., R. Greenberg, and R. Barnes
2008. Tidal Evolution of Close-in Extrasolar Planets. *ApJ*, 678(2):1396–1406.

Jacob, W. S.
1855. On certain Anomalies presented by the Binary Star 70 Ophiuchi. *MNRAS*, 15:228.

Jenkins, J. M., J. D. Twicken, N. M. Batalha, D. A. Caldwell, W. D. Cochran, M. Endl, D. W. Latham, G. A. Esquerdo, S. Seader, A. Bieryla, E. Petigura, D. R. Ciardi, G. W. Marcy, H. Isaacson, D. Huber, J. F. Rowe, G. Torres, S. T. Bryson, L. Buchhave, I. Ramirez, A. Wolfgang, J. Li, J. R. Campbell, P. Tenenbaum, D. Sand erfer, C. E. Henze, J. H. Catanzarite, R. L. Gilliland, and W. J. Borucki
2015. Discovery and Validation of Kepler-452b: A 1.6 R $_?$ Super Earth Exoplanet in the Habitable Zone of a G2 Star. *AJ*, 150(2):56.

Jin, S. and C. Mordasini
2018. Compositional Imprints in Density-Distance-Time: A Rocky Composition for Close-in Low-mass Exoplanets from the Location of the Valley of Evaporation. *ApJ*, 853(2):163.

Johansen, A., M. B. Davies, R. P. Church, and V. Holmelin
2012. Can Planetary Instability Explain the Kepler Dichotomy? *ApJ*, 758:39.

Johansen, A. and P. Lacerda
2010. Prograde rotation of protoplanets by accretion of pebbles in a gaseous environment. *MNRAS*, 404(1):475–485.

Johansen, A. and M. Lambrechts
2017. Forming Planets via Pebble Accretion. *Annual Review of Earth and Planetary Sciences*, 45(1):359–387.

Johansen, A., J. S. Oishi, M.-M. Mac Low, H. Klahr, T. Henning, and A. Youdin
2007. Rapid planetesimal formation in turbulent circumstellar disks. *Nature*, 448(7157):1022–1025.

Jurić, M. and S. Tremaine
2008. Dynamical Origin of Extrasolar Planet Eccentricity Distribution. *ApJ*, 686(1):603–620.

Kane, M., D. Ragozzine, X. Flowers, T. Holczer, T. Mazeh, and H. M. Relles
2019. Visual Analysis and Demographics of Kepler Transit Timing Variations. *AJ*, 157(5):171.

Kant, I.
1755. *Allgemeine Naturgeschichte und Theorie des Himmels*.

Kiefer, F.
2019. Determining the mass of the planetary candidate HD 114762 b using Gaia. *A&A*, 632:L9.

Kiefer, F., G. Hébrard, J. Sahlmann, S. G. Sousa, T. Forveille, N. Santos, M. Mayor, M. Deleuil, P. A. Wilson, S. Dalal, R. F. Díaz, G. W. Henry, J. Hagelberg, M. J. Hobson, O. Demangeon, V. Bourrier, X. Delfosse, L. Arnold, N. Astudillo-Defru, J. L. Beuzit, I. Boisse, X. Bonfils, S. Borgniet, F. Bouchy, B. Courcol, D. Ehrenreich, N. Hara, A. M. Lagrange, C. Lovis, G. Montagnier, C. Moutou, F. Pepe, C. Perrier, J. Rey, A. Santerne, D. Ségransan, S. Udry, and A. Vidal-Madjar
2019. Detection and characterisation of 54 massive companions with the SOPHIE spectrograph. Seven new brown dwarfs and constraints on the brown dwarf desert. *A&A*, 631:A125.

Kley, W. and R. P. Nelson
2012. Planet-Disk Interaction and Orbital Evolution. *ARAA*, 50:211–249.

Koch, D. G., W. J. Borucki, J. F. Rowe, N. M. Batalha, T. M. Brown, D. A. Caldwell, J. Caldwell, W. D. Cochran, E. DeVore, E. W. Dunham, A. K. Dupree, I. Gautier, Thomas N., J. C. Geary, R. L. Gilliland, S. B. Howell, J. M. Jenkins, D. W. Latham, J. J. Lissauer, G. W. Marcy, D. Morrison, and J. Tarter
2010. Discovery of the Transiting Planet Kepler-5b. *ApJL*, 713(2):L131–L135.

Kokubo, E.
2005. Dynamics of planetesimals: the role of two-body relaxation. In *IAU Colloq. 197: Dynamics of Populations of Planetary Systems*, Z. Knežević and A. Milani, eds., Pp. 41–46.

Kokubo, E. and H. Genda
2010. Formation of Terrestrial Planets from Protoplanets Under a Realistic Accretion Condition. *ApJL*, 714(1):L21–L25.

Kokubo, E. and S. Ida
1998. Oligarchic Growth of Protoplanets. *Icarus*, 131(1):171–178.

Lai, D. and B. Pu
2017. Hiding Planets behind a Big Friend: Mutual Inclinations of Multi-planet Systems with External Companions. *AJ*, 153(1):42.

Lambrechts, M. and A. Johansen
2012. Rapid growth of gas-giant cores by pebble accretion. *A&A*, 544:A32.

Lambrechts, M., A. Morbidelli, S. A. Jacobson, A. Johansen, B. Bitsch, A. Izidoro, and S. N. Raymond
2019. Formation of planetary systems by pebble accretion and migration. How the radial pebble flux determines a terrestrial-planet or super-Earth growth mode. *A&A*, 627:A83.

Laskar, J. and A. C. Petit
2017. AMD-stability and the classification of planetary systems. *A&A*, 605:A72.

Latham, D. W., R. P. Stefanik, T. Mazeh, M. Mayor, and G. Burki
1989. The unseen companion of HD114762 - A probable brown dwarf. *Nature*, 339:38–40.

Lee, E. J., E. Chiang, and C. W. Ormel
2014. Make Super-Earths, Not Jupiters: Accreting Nebular Gas onto Solid Cores at 0.1 AU and Beyond. *ApJ*, 797(2):95.

Leinhardt, Z. M. and S. T. Stewart
2012. Collisions between Gravity-dominated Bodies. I. Outcome Regimes and Scaling Laws. *ApJ*, 745:79.

Levison, H. F. and M. J. Duncan
1994. The Long-Term Dynamical Behavior of Short-Period Comets. *Icarus*, 108:18–36.

Lin, D. N. C., P. Bodenheimer, and D. C. Richardson
1996. Orbital migration of the planetary companion of 51 Pegasi to its present location. *Nature*, 380(6575):606–607.

Lin, D. N. C. and J. Papaloizou
1986. On the Tidal Interaction between Protoplanets and the Protoplanetary Disk. III. Orbital Migration of Protoplanets. *ApJ*, 309:846.

Lissauer, J. J., D. C. Fabrycky, E. B. Ford, W. J. Borucki, F. Fressin, G. W. Marcy, J. A. Orosz, J. F. Rowe, G. Torres, W. F. Welsh, N. M. Batalha, S. T. Bryson, L. A. Buchhave, D. A. Caldwell, J. A. Carter, D. Charbonneau, J. L. Christiansen, W. D. Cochran, J.-M. Desert, E. W. Dunham, M. N. Fanelli, J. J. Fortney, I. Gautier, Thomas N., J. C. Geary, R. L. Gilliland, M. R. Haas, J. R. Hall, M. J. Holman, D. G. Koch, D. W. Latham, E. Lopez, S. McCauliff, N. Miller, R. C. Morehead, E. V. Quintana, D. Ragozzine, D. Sasselov, D. R. Short, and J. H. Steffen
2011a. A closely packed system of low-mass, low-density planets transiting Kepler-11. *Nature*, 470(7332):53–58.

Lissauer, J. J., D. Jontof-Hutter, J. F. Rowe, D. C. Fabrycky, E. D. Lopez, E. Agol, G. W. Marcy, K. M. Deck, D. A. Fischer, J. J. Fortney, S. B. Howell, H. Isaacson, J. M. Jenkins, R. Kolbl, D. Sasselov, D. R. Short, and W. F. Welsh
2013. All Six Planets Known to Orbit Kepler-11 Have Low Densities. *ApJ*, 770(2):131.

Lissauer, J. J., D. Ragozzine, D. C. Fabrycky, J. H. Steffen, E. B. Ford, J. M. Jenkins, A. Shporer, M. J. Holman, J. F. Rowe, E. V. Quintana, N. M. Batalha, W. J. Borucki, S. T. Bryson, D. A. Caldwell, J. A. Carter, D. Ciardi, E. W. Dunham, J. J. Fortney, I. Gautier, Thomas N., S. B. Howell, D. G. Koch, D. W. Latham, G. W. Marcy, R. C. Morehead, and D. Sasselov
2011b. Architecture and Dynamics of Kepler's Candidate Multiple Transiting Planet Systems. *ApJS*, 197:8.

Liu, B. and J. Ji
2020. A tale of planet formation: from dust to planets. *Research in Astronomy and Astrophysics*, 20(10):164.

Lopez, E. D. and J. J. Fortney
2014. Understanding the Mass-Radius Relation for Sub-neptunes: Radius as a Proxy for Composition. *ApJ*, 792(1):1.

Lopez, E. D., J. J. Fortney, and N. Miller
2012. How Thermal Evolution and Mass-loss Sculpt Populations of Super-Earths and Sub-Neptunes: Application to the Kepler-11 System and Beyond. *ApJ*, 761(1):59.

Lovis, C. and D. Fischer
2010. *Radial Velocity Techniques for Exoplanets*, Pp. 27–53.

Lovis, C., D. Ségransan, M. Mayor, S. Udry, W. Benz, J. L. Bertaux, F. Bouchy, A. C. M. Correia, J. Laskar, G. Lo Curto, C. Mordasini, F. Pepe, D. Queloz, and N. C. Santos
2011. The HARPS search for southern extra-solar planets. XXVIII. Up to seven planets orbiting HD 10180: probing the architecture of low-mass planetary systems. *A&A*, 528:A112.

MacDonald, M. G., R. I. Dawson, S. J. Morrison, E. J. Lee, and A. Khandelwal
2020. Forming Diverse Super-Earth Systems In Situ. *ApJ*, 891(1):20.

MacDonald, M. G., D. Ragozzine, D. C. Fabrycky, E. B. Ford, M. J. Holman, H. T. Isaacson, J. J. Lissauer, E. D. Lopez, T. Mazeh, L. Rogers, J. F. Rowe, J. H. Steffen, and G. Torres
2016. A Dynamical Analysis of the Kepler-80 System of Five Transiting Planets. *AJ*, 152:105.

Mahajan, N. and Y. Wu
2014. Stability of the Kepler-11 System and Its Origin. *ApJ*, 795(1):32.

Mandell, A. M., S. N. Raymond, and S. Sigurdsson
2007. Formation of Earth-like Planets During and After Giant Planet Migration. *ApJ*, 660(1):823–844.

Marcus, R. A., D. Sasselov, S. T. Stewart, and L. Hernquist
2010. Water/Icy Super-Earths: Giant Impacts and Maximum Water Content. *ApJ*, 719:L45–L49.

Marcy, G. W., H. Isaacson, A. W. Howard, J. F. Rowe, J. M. Jenkins, S. T. Bryson, D. W. Latham, S. B. Howell, I. Gautier, Thomas N., N. M. Batalha, L. Rogers, D. Ciardi, D. A. Fischer, R. L. Gilliland, H. Kjeldsen, J. Christensen-Dalsgaard, D. Huber, W. J. Chaplin, S. Basu, L. A. Buchhave, S. N. Quinn, W. J. Borucki, D. G. Koch, R. Hunter, D. A. Caldwell, J. Van Cleve, R. Kolbl, L. M. Weiss, E. Petigura, S. Seager, T. Morton, J. A. Johnson, S. Ballard, C. Burke, W. D. Cochran, M. Endl, P. MacQueen, M. E. Everett, J. J. Lissauer, E. B. Ford, G. Torres, F. Fressin, T. M. Brown, J. H. Steffen, D. Charbonneau, G. S. Basri, D. D. Sasselov, J. Winn, R. Sanchis-Ojeda, J. Christiansen, E. Adams, C. Henze, A. Dupree, D. C. Fabrycky, J. J. Fortney, J. Tarter, M. J. Holman, P. Tenenbaum, A. Shporer, P. W. Lucas, W. F. Welsh, J. A. Orosz, T. R. Bedding, T. L. Campante, G. R. Davies, Y. Elsworth, R. Handberg, S. Hekker, C. Karoff, S. D. Kawaler, M. N. Lund, M. Lundkvist, T. S. Metcalfe, A. Miglio, V. Silva Aguirre, D. Stello, T. R. White, A. Boss, E. Devore, A. Gould, A. Prsa, E. Agol, T. Barclay, J. Coughlin, E. Brugamyer, F. Mullally, E. V. Quintana, M. Still, S. E. Thompson, D. Morrison, J. D. Twicken, J.-M. Désert, J. Carter, J. R. Crepp, G. Hébrard, A. Santerne, C. Moutou, C. Sobeck, D. Hudgins, M. R. Haas, P. Robertson, J. Lillo-Box, and D. Barrado
2014. Masses, Radii, and Orbits of Small Kepler Planets: The Transition from Gaseous to Rocky Planets. *ApJS*, 210(2):20.

Marois, C., B. Macintosh, T. Barman, B. Zuckerman, I. Song, J. Patience, D. Lafrenière, and R. Doyon
2008. Direct Imaging of Multiple Planets Orbiting the Star HR 8799. *Science*, 322(5906):1348.

Marzari, F. and M. Nagasawa
2020. Secular evolution of close-in planets: the effects of general relativity. *MNRAS*, 493(1):427–436.

Masuda, K., J. N. Winn, and H. Kawahara
2020. Mutual Orbital Inclinations between Cold Jupiters and Inner Super-Earths. *AJ*, 159(2):38.

Matsumoto, Y., P.-G. Gu, E. Kokubo, S. Oshino, and M. Omiya
2020. Ejection of close-in super-Earths around low-mass stars in the giant impact stage. *A&A*, 642:A23.

Matsumoto, Y. and E. Kokubo
2017. Formation of Close-in Super-Earths by Giant Impacts: Effects of Initial Eccentricities and Inclinations of Protoplanets. *AJ*, 154:27.

Matsumoto, Y., M. Nagasawa, and S. Ida
2012. The orbital stability of planets trapped in the first-order mean-motion resonances. *Icarus*, 221(2):624–631.

Matsumura, S., S. Ida, and M. Nagasawa
2013. Effects of Dynamical Evolution of Giant Planets on Survival of Terrestrial Planets. *ApJ*, 767(2):129.

Matsuyama, I., D. Johnstone, and L. Hartmann
2003. Viscous Diffusion and Photoevaporation of Stellar Disks. *ApJ*, 582(2):893–904.

Mayor, M., M. Marmier, C. Lovis, D. Udry, S. a2019MNRASnd Ségransan, F. Pepe, W. Benz, J. L. Bertaux, F. Bouchy, X. Dumusque, G. Lo Curto, C. Mordasini, D. Queloz, and N. C. Santos
2011. The HARPS search for southern extra-solar planets XXXIV. Occurrence, mass distribution and orbital properties of super-Earths and Neptune-mass planets. *arXiv e-prints*, P. arXiv:1109.2497.

Mayor, M., F. Pepe, D. Queloz, F. Bouchy, G. Rupprecht, G. Lo Curto, G. Avila, W. Benz, J.-L. Bertaux, X. Bonfils, T. Dall, H. Dekker, B. Delabre, W. Eckert, M. Fleury, A. Gilliotte, D. Gojak, J. C. Guzman, D. Kohler, J.-L. Lizon, A. Longinotti, C. Lovis, D. Megevand, L. Pasquini, J. Reyes, J.-P. Sivan, D. Sosnowska, R. Soto, S. Udry, A. van Kesteren, L. Weber, and U. Weilenmann
2003. Setting New Standards with HARPS. *The Messenger*, 114:20–24.

Mayor, M. and D. Queloz
1995. A Jupiter-mass companion to a solar-type star. *Nature*, 378:355–359.

McNally, C. P., R. P. Nelson, and S.-J. Paardekooper
2019a. Multiplanet systems in inviscid discs can avoid forming resonant chains. *MNRAS*, 489(1):L17–L21.

McNally, C. P., R. P. Nelson, S.-J. Paardekooper, and P. Benítez-Llambay
2019b. Migrating super-Earths in low-viscosity discs: unveiling the roles of feedback, vortices, and laminar accretion flows. *MNRAS*, 484(1):728–748.

Mills, S. M., A. W. Howard, E. A. Petigura, B. J. Fulton, H. Isaacson, and L. M. Weiss
2019. The California-Kepler Survey. VIII. Eccentricities of Kepler Planets and Tentative Evidence of a High-metallicity Preference for Small Eccentric Planets. *AJ*, 157(5):198.

Mizuno, H., K. Nakazawa, and C. Hayashi
1978. Instability of a gaseous envelope surrounding a planetary core and formation of giant planets. *Progress of Theoretical Physics*, 60:699–710.

Moorhead, A. V., E. B. Ford, R. C. Morehead, J. Rowe, W. J. Borucki, N. M. Batalha, S. T. Bryson, D. A. Caldwell, D. C. Fabrycky, I. Gautier, Thomas N.,

D. G. Koch, M. J. Holman, J. M. Jenkins, J. Li, J. J. Lissauer, P. Lucas, G. W. Marcy, S. N. Quinn, E. Quintana, D. Ragozzine, A. Shporer, M. Still, and G. Torres
2011. The Distribution of Transit Durations for Kepler Planet Candidates and Implications for Their Orbital Eccentricities. *ApJS*, 197(1):1.

Morbidelli, A. and D. Nesvorny
2012. Dynamics of pebbles in the vicinity of a growing planetary embryo: hydrodynamical simulations. *A&A*, 546:A18.

Mordasini, C., H. Klahr, Y. Alibert, N. Miller, and T. Henning
2014. Grain opacity and the bulk composition of extrasolar planets. I. Results from scaling the ISM opacity. *A&A*, 566:A141.

Moriarty, J. and S. Ballard
2016. The Kepler Dichotomy in Planetary Disks: Linking Kepler Observables to Simulations of Late-stage Planet Formation. *ApJ*, 832(1):34.

Morton, T. D., S. T. Bryson, J. L. Coughlin, J. F. Rowe, G. Ravichandran, E. A. Petigura, M. R. Haas, and N. M. Batalha
2016. False Positive Probabilities for all Kepler Objects of Interest: 1284 Newly Validated Planets and 428 Likely False Positives. *ApJ*, 822:86.

Motalebi, F., S. Udry, M. Gillon, C. Lovis, D. Ségransan, L. A. Buchhave, B. O. Demory, L. Malavolta, C. D. Dressing, D. Sasselov, K. Rice, D. Charbonneau, A. Collier Cameron, D. Latham, E. Molinari, F. Pepe, L. Affer, A. S. Bonomo, R. Cosentino, X. Dumusque, P. Figueira, A. F. M. Fiorenzano, S. Gettel, A. Harutyunyan, R. D. Haywood, J. Johnson, E. Lopez, M. Lopez-Morales, M. Mayor, G. Micela, A. Mortier, V. Nascimbeni, D. Philips, G. Piotto, D. Pollacco, D. Queloz, A. Sozzetti, A. Vand erburg, and C. A. Watson
2015. The HARPS-N Rocky Planet Search. I. HD 219134 b: A transiting rocky planet in a multi-planet system at 6.5 pc from the Sun. *A&A*, 584:A72.

Movshovitz, N., P. Bodenheimer, M. Podolak, and J. J. Lissauer
2010. Formation of Jupiter using opacities based on detailed grain physics. *Icarus*, 209(2):616–624.

Mullally, F., J. L. Coughlin, S. E. Thompson, J. Rowe, C. Burke, D. W. Latham, N. M. Batalha, S. T. Bryson, J. Christiansen, C. E. Henze, A. Ofir, B. Quarles, A. Shporer, V. Van Eylen, C. Van Laerhoven, Y. Shah, A. Wolfgang, W. J. Chaplin, J.-W. Xie, R. Akeson, V. Argabright, E. Bachtell, T. Barclay, W. J. Borucki, D. A. Caldwell, J. R. Campbell, J. H. Catanzarite, W. D. Cochran, R. M. Duren, S. W. Fleming, D. Fraquelli, F. R. Girouard, M. R. Haas, K. G. Hełminiak, S. B. Howell, D. Huber, K. Larson, I. Gautier, Thomas N., J. M. Jenkins, J. Li, J. J. Lissauer, S. McArthur, C. Miller, R. L. Morris, A. Patil-Sabale, P. Plavchan, D. Putnam, E. V. Quintana, S. Ramirez, V. Silva Aguirre, S. Seader, J. C. Smith, J. H. Steffen, C. Stewart, J. Stober, M. Still, P. Tenenbaum, J. Troeltzsch, J. D.

Twicken, and K. A. Zamudio
2015. Planetary Candidates Observed by Kepler. VI. Planet Sample from Q1–Q16 (47 Months). *The Astrophysical Journal Supplement Series*, 217:31.

Murray, C. D. and S. F. Dermott
1999. *Solar system dynamics*. Cambridge University Press.

Mustill, A. J., M. B. Davies, and A. Johansen
2015. The Destruction of Inner Planetary Systems during High-eccentricity Migration of Gas Giants. *ApJ*, 808(1):14.

Mustill, A. J., M. B. Davies, and A. Johansen
2017. The effects of external planets on inner systems: multiplicities, inclinations and pathways to eccentric warm Jupiters. *MNRAS*, 468(3):3000–3023.

Mustill, A. J., M. B. Davies, and A. Johansen
2018. The dynamical evolution of transiting planetary systems including a realistic collision prescription. *MNRAS*, 478:2896–2908.

Nelson, R. P. and J. C. B. Papaloizou
2004. The interaction of giant planets with a disc with MHD turbulence - IV. Migration rates of embedded protoplanets. *MNRAS*, 350(3):849–864.

Nettelmann, N., J. J. Fortney, U. Kramm, and R. Redmer
2011. Thermal Evolution and Structure Models of the Transiting Super-Earth GJ 1214b. *ApJ*, 733(1):2.

Ogihara, M., A. Morbidelli, and T. Guillot
2015. A reassessment of the in situ formation of close-in super-Earths. *A&A*, 578:A36.

Ormel, C. W. and H. H. Klahr
2010. The effect of gas drag on the growth of protoplanets. Analytical expressions for the accretion of small bodies in laminar disks. *A&A*, 520:A43.

Owen, J. E. and B. Campos Estrada
2020. Testing exoplanet evaporation with multitransiting systems. *MNRAS*, 491(4):5287–5297.

Owen, J. E. and Y. Wu
2013. Kepler Planets: A Tale of Evaporation. *ApJ*, 775(2):105.

Owen, J. E. and Y. Wu
2017. The Evaporation Valley in the Kepler Planets. *ApJ*, 847(1):29.

Papaloizou, J. C. B. and J. D. Larwood
2000. On the orbital evolution and growth of protoplanets embedded in a gaseous disc. *MNRAS*, 315(4):823–833.

Papaloizou, J. C. B. and R. P. Nelson
2005. Models of accreting gas giant protoplanets in protostellar disks. *A&A*, 433(1):247–265.

Papaloizou, J. C. B. and C. Terquem
1999. Critical Protoplanetary Core Masses in Protoplanetary Disks and the Formation of Short-Period Giant Planets. *ApJ*, 521(2):823–838.

Pepe, F., S. Cristiani, R. Rebolo, N. C. Santos, H. Dekker, A. Cabral, P. Di Marcantonio, P. Figueira, G. Lo Curto, C. Lovis, M. Mayor, D. Mégevand, P. Molaro, M. Riva, M. R. Zapatero Osorio, M. Amate, A. Manescau, L. Pasquini, F. M. Zerbi, V. Adibekyan, M. Abreu, M. Affolter, Y. Alibert, M. Aliverti, R. Allart, C. Allende Prieto, D. Álvarez, D. Alves, G. Avila, V. Baldini, T. Bandy, S. C. C. Barros, W. Benz, A. Bianco, F. Borsa, V. Bourrier, F. Bouchy, C. Broeg, G. Calderone, R. Cirami, J. Coelho, P. Conconi, I. Coretti, C. Cumani, G. Cupani, V. D'Odorico, M. Damasso, S. Deiries, B. Delabre, O. D. S. Demangeon, X. Dumusque, D. Ehrenreich, J. P. Faria, A. Fragoso, L. Genolet, M. Genoni, R. Génova Santos, J. I. González Hernández, I. Hughes, O. Iwert, F. Kerber, J. Knudstrup, M. Landoni, B. Lavie, J. Lillo-Box, J. L. Lizon, C. Maire, C. J. . A. P. Martins, A. Mehner, G. Micela, A. Modigliani, M. A. Monteiro, M. J. P. F. G. Monteiro, M. Moschetti, M. T. Murphy, N. Nunes, L. Oggioni, A. Oliveira, M. Oshagh, E. Pallé, G. Pariani, E. Poretti, J. L. Rasilla, J. Rebordão, E. M. Redaelli, S. Santana Tschudi, P. Santin, P. Santos, D. Ségransan, T. M. Schmidt, A. Segovia, D. Sosnowska, A. Sozzetti, S. G. Sousa, P. Spanò, A. Suárez Mascareño, H. Tabernero, F. Tenegi, S. Udry, and A. Zanutta
2020. ESPRESSO@VLT – On-sky performance and first results. *arXiv e-prints*, P. arXiv:2010.00316.

Perri, F. and A. G. W. Cameron
1974. Hydrodynamic Instability of the Solar Nebula in the Presence of a Planetary Core. *Icarus*, 22(4):416–425.

Perryman, M., J. Hartman, G. Á. Bakos, and L. Lindegren
2014. Astrometric Exoplanet Detection with Gaia. *ApJ*, 797(1):14.

Perryman, M. A. C., K. S. de Boer, G. Gilmore, E. Høg, M. G. Lattanzi, L. Lindegren, X. Luri, F. Mignard, O. Pace, and P. T. de Zeeuw
2001. GAIA: Composition, formation and evolution of the Galaxy. *A&A*, 369:339–363.

Petigura, E. A., A. W. Howard, and G. W. Marcy
2013. Prevalence of Earth-size planets orbiting Sun-like stars. *Proceedings of the National Academy of Science*, 110(48):19273–19278.

Petrovich, C., R. Malhotra, and S. Tremaine
2013. Planets near Mean-motion Resonances. *ApJ*, 770:24.

Petrovich, C., S. Tremaine, and R. Rafikov
2014. Scattering Outcomes of Close-in Planets: Constraints on Planet Migration. *ApJ*, 786(2):101.

Pinte, C., G. van der Plas, F. Ménard, D. J. Price, V. Christiaens, T. Hill, D. Mentiplay, C. Ginski, E. Choquet, Y. Boehler, G. Duchêne, S. Perez, and S. Casassus
2019. Kinematic detection of a planet carving a gap in a protoplanetary disk. *Nature Astronomy*, 3:1109–1114.

Piso, A.-M. A. and A. N. Youdin
2014. On the Minimum Core Mass for Giant Planet Formation at Wide Separations. *ApJ*, 786(1):21.

Pollacco, D. L., I. Skillen, A. Collier Cameron, D. J. Christian, C. Hellier, J. Irwin, T. A. Lister, R. A. Street, R. G. West, D. R. Anderson, W. I. Clarkson, H. Deeg, B. Enoch, A. Evans, A. Fitzsimmons, C. A. Haswell, S. Hodgkin, K. Horne, S. R. Kane, F. P. Keenan, P. F. L. Maxted, A. J. Norton, J. Osborne, N. R. Parley, R. S. I. Ryans, B. Smalley, P. J. Wheatley, and D. M. Wilson
2006. The WASP Project and the SuperWASP Cameras. *PASP*, 118(848):1407–1418.

Pollack, J. B., O. Hubickyj, P. Bodenheimer, J. J. Lissauer, M. Podolak, and Y. Greenzweig
1996. Formation of the Giant Planets by Concurrent Accretion of Solids and Gas. *Icarus*, 124(1):62–85.

Poon, S. T. S.
2016. *Intrinsic Inclinations of Kepler Planetary System*. M.Sc. Thesis, Queen Mary University of London.

Poon, S. T. S. and R. P. Nelson
2020. On the origin of the eccentricity dichotomy displayed by compact super-Earths: dynamical heating by cold giants. *MNRAS*, 498(4):5166–5182.

Poon, S. T. S., R. P. Nelson, and G. A. L. Coleman
. In situ formation of hot jupiters with companion super-earths. submitted.

Poon, S. T. S., R. P. Nelson, S. A. Jacobson, and A. Morbidelli
2020. Formation of compact systems of super-Earths via dynamical instabilities and giant impacts. *MNRAS*, 491(4):5595–5620.

Poynting, J. H.
1903. Radiation in the solar system : its effect on temperature and its pressure on small bodies. *MNRAS*, 64:1.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery
1992. *Numerical recipes in FORTRAN. The art of scientific computing*. Cambridge: University Press, |c1992, 2nd ed.

Pringle, J. E.
1981. Accretion discs in astrophysics. *ARAA*, 19:137–162.

Pu, B. and D. Lai
2018. Eccentricities and inclinations of multiplanet systems with external perturbers. *MNRAS*, 478(1):197–217.

Pu, B. and Y. Wu
2015. Spacing of Kepler Planets: Sculpting by Dynamical Instability. *ApJ*, 807:44.

Quinn, S. N., R. J. White, D. W. Latham, L. A. Buchhave, J. R. Cantrell, S. E. Dahm, G. Fűrész, A. H. Szentgyorgyi, J. C. Geary, G. Torres, A. Bieryla, P. Berlind, M. C. Calkins, G. A. Esquerdo, and R. P. Stefanik
2012. Two "b"s in the Beehive: The Discovery of the First Hot Jupiters in an Open Cluster. *ApJL*, 756(2):L33.

Rauer, H., C. Catala, C. Aerts, T. Appourchaux, W. Benz, A. Brandeker, J. Christensen-Dalsgaard, M. Deleuil, L. Gizon, M. J. Goupil, M. Güdel, E. Janot-Pacheco, M. Mas-Hesse, I. Pagano, G. Piotto, D. Pollacco, Ċ. Santos, A. Smith, J. C. Suárez, R. Szabó, S. Udry, V. Adibekyan, Y. Alibert, J. M. Almenara, P. Amaro-Seoane, M. A.-v. Eiff, M. Asplund, E. Antonello, S. Barnes, F. Baudin, K. Belkacem, M. Bergemann, G. Bihain, A. C. Birch, X. Bonfils, I. Boisse, A. S. Bonomo, F. Borsa, I. M. Brand ão, E. Brocato, S. Brun, M. Burleigh, R. Burston, J. Cabrera, S. Cassisi, W. Chaplin, S. Charpinet, C. Chiappini, R. P. Church, S. Csizmadia, M. Cunha, M. Damasso, M. B. Davies, H. J. Deeg, R. F. Díaz, S. Dreizler, C. Dreyer, P. Eggenberger, D. Ehrenreich, P. Eigmüller, A. Erikson, R. Farmer, S. Feltzing, F. de Oliveira Fialho, P. Figueira, T. Forveille, M. Fridlund, R. A. García, P. Giommi, G. Giuffrida, M. Godolt, J. Gomes da Silva, T. Granzer, J. L. Grenfell, A. Grotsch-Noels, E. Günther, C. A. Haswell, A. P. Hatzes, G. Hébrard, S. Hekker, R. Helled, K. Heng, J. M. Jenkins, A. Johansen, M. L. Khodachenko, K. G. Kislyakova, W. Kley, U. Kolb, N. Krivova, F. Kupka, H. Lammer, A. F. Lanza, Y. Lebreton, D. Magrin, P. Marcos-Arenal, P. M. Marrese, J. P. Marques, J. Martins, S. Mathis, S. Mathur, S. Messina, A. Miglio, J. Montalban, M. Montalto, M. J. P. F. G. Monteiro, H. Moradi, E. Moravveji, C. Mordasini, T. Morel, A. Mortier, V. Nascimbeni, R. P. Nelson, M. B. Nielsen, L. Noack, A. J. Norton, A. Ofir, M. Oshagh, R. M. Ouazzani, P. Pápics, V. C. Parro, P. Petit, B. Plez, E. Poretti, A. Quirrenbach, R. Ragazzoni, G. Raimondo, M. Rainer, D. R. Reese, R. Redmer, S. Reffert, B. Rojas-Ayala, I. W. Roxburgh, S. Salmon, A. Santerne, J. Schneider, J. Schou, S. Schuh, H. Schunker, A. Silva-Valio, R. Silvotti, I. Skillen, I. Snellen, F. Sohl, S. G. Sousa, A. Sozzetti, D. Stello, K. G. Strassmeier, M. Švanda, G. M. Szabó, A. Tkachenko, D. Valencia, V. Van Grootel, S. D. Vauclair, P. Ventura, F. W. Wagner, N. A. Walton, J. Weingrill, S. C. Werner, P. J. Wheatley, and K. Zwintz
2014. The PLATO 2.0 mission. *Experimental Astronomy*, 38(1-2):249–330.

Raymond, S. N., A. M. Mandell, and S. Sigurdsson
2006. Exotic Earths: Forming Habitable Worlds with Giant Planet Migration. *Science*, 313(5792):1413–1416.

Ribas, I., M. Tuomi, A. Reiners, R. P. Butler, J. C. Morales, M. Perger, S. Dreizler, C. Rodríguez-López, J. I. González Hernández, A. Rosich, F. Feng, T. Trifonov, S. S. Vogt, J. A. Caballero, A. Hatzes, E. Herrero, S. V. Jeffers, M. Lafarga, F. Murgas, R. P. Nelson, E. Rodríguez, J. B. P. Strachan, L. Tal-Or, J. Teske, B. Toledo-Padrón, M. Zechmeister, A. Quirrenbach, P. J. Amado, M. Azzaro, V. J. S. Béjar, J. R. Barnes, Z. M. Berdiñas, J. Burt, G. Coleman, M. Cortés-Contreras, J. Crane, S. G. Engle, E. F. Guinan, C. A. Haswell, T. Henning, B. Holden, J. Jenkins, H. R. A. Jones, A. Kaminski, M. Kiraga, M. Kürster, M. H. Lee, M. J. López-González, D. Montes, J. Morin, A. Ofir, E. Pallé, R. Rebolo, S. Reffert, A. Schweitzer, W. Seifert, S. A. Shectman, D. Staab, R. A. Street, A. Suárez Mascareño, Y. Tsapras, S. X. Wang, and G. Anglada-Escudé
2018. A candidate super-Earth planet orbiting near the snow line of Barnard's star. *Nature*, 563(7731):365–368.

Rivera, E. J., J. J. Lissauer, R. P. Butler, G. W. Marcy, S. S. Vogt, D. A. Fischer, T. M. Brown, G. Laughlin, and G. W. Henry
2005. A ~7.5 $M_\oplus$ Planet Orbiting the Nearby Star, GJ 876. *ApJ*, 634(1):625–640.

Robertson, H. P.
1937. Dynamical effects of radiation in the solar system. *MNRAS*, 97:423.

Rowe, J. F., S. T. Bryson, G. W. Marcy, J. J. Lissauer, D. Jontof-Hutter, F. Mullally, R. L. Gilliland, H. Issacson, E. Ford, S. B. Howell, W. J. Borucki, M. Haas, D. Huber, J. H. Steffen, S. E. Thompson, E. Quintana, T. Barclay, M. Still, J. Fortney, T. N. Gautier, III, R. Hunter, D. A. Caldwell, D. R. Ciardi, E. Devore, W. Cochran, J. Jenkins, E. Agol, J. A. Carter, and J. Geary
2014. Validation of Kepler's Multiple Planet Candidates. III. Light Curve Analysis and Announcement of Hundreds of New Multi-planet Systems. *ApJ*, 784:45.

Rowe, J. F., J. L. Coughlin, V. Antoci, T. Barclay, N. M. Batalha, W. J. Borucki, C. J. Burke, S. T. Bryson, D. A. Caldwell, J. R. Campbell, J. H. Catanzarite, J. L. Christiansen, W. Cochran, R. L. Gilliland, F. R. Girouard, M. R. Haas, K. G. Hełminiak, C. E. Henze, K. L. Hoffman, S. B. Howell, D. Huber, R. C. Hunter, H. Jang-Condell, J. M. Jenkins, T. C. Klaus, D. W. Latham, J. Li, J. J. Lissauer, S. D. McCauliff, R. L. Morris, F. Mullally, A. Ofir, B. Quarles, E. Quintana, A. Sabale, S. Seader, A. Shporer, J. C. Smith, J. H. Steffen, M. Still, P. Tenenbaum, S. E. Thompson, J. D. Twicken, C. Van Laerhoven, A. Wolfgang, and K. A. Zamudio
2015. Planetary Candidates Observed by Kepler. V. Planet Sample from Q1-Q12 (36 Months). *The Astrophysical Journal Supplement Series*, 217:16.

Ruden, S. P.
2004. Evolution of Photoevaporating Protoplanetary Disks. *ApJ*, 605(2):880–891.

Safronov, V. S.
1972. *Evolution of the protoplanetary cloud and formation of the earth and planets.*

Sahlmann, J., P. F. Lazorenko, D. Ségransan, E. L. Martín, D. Queloz, M. Mayor, and S. Udry
2013. Astrometric orbit of a low-mass companion to an ultracool dwarf. *A&A*, 556:A133.

Santerne, A., R. F. Díaz, C. Moutou, F. Bouchy, G. Hébrard, J.-M. Almenara, A. S. Bonomo, M. Deleuil, and N. C. Santos
2012. SOPHIE velocimetry of Kepler transit candidates. VII. A false-positive rate of 35% for Kepler close-in giant candidates. *A&A*, 545:A76.

Schmitt, J. R., J. M. Jenkins, and D. A. Fischer
2017. A Search for Lost Planets in the Kepler Multi-planet Systems and the Discovery of the Long-period, Neptune-sized Exoplanet Kepler-150 f. *AJ*, 153(4):180.

Schöck, M., S. Els, R. Riddle, W. Skidmore, T. Travouillon, R. Blum, E. Bustos, G. Chanan, S. G. Djorgovski, P. Gillett, B. Gregory, J. Nelson, A. Otárola, J. Seguel, J. Vasquez, A. Walker, D. Walker, and L. Wang
2009. Thirty Meter Telescope Site Testing I: Overview. *PASP*, 121(878):384.

Scora, J., D. Valencia, A. Morbidelli, and S. Jacobson
2020. Chemical diversity of super-Earths as a consequence of formation. *MNRAS*, 493(4):4910–4924.

Seager, S.
2010. *Exoplanets*. University of Arizona Press.

See, T. J. J.
1896. Researches on the orbit of 70 Ophiuchi, and on a periodic perturbation in the motion of the system arising from the action of an unseen body. *AJ*, 16:17–23.

Shakura, N. I. and R. A. Sunyaev
1973. Reprint of 1973A&A....24..337S. Black holes in binary systems. Observational appearance. *A&A*, 500:33–51.

Shallue, C. J. and A. Vanderburg
2018. Identifying Exoplanets with Deep Learning: A Five-planet Resonant Chain around Kepler-80 and an Eighth Planet around Kepler-90. *AJ*, 155:94.

Sherrill, T. J.
1999. A Career of Controversy: The Anomaly of T. J. J. See. *Journal for the History of Astronomy*, 30:25.

Shoemaker, E. M.
1962. *Interpretation of lunar crater*. Physics and Astronomy of the Moon (NewYork: Academic).

Shvartzvald, Y., J. C. Yee, S. Calchi Novati, A. Gould, C. U. Lee, C. Beichman, G. Bryden, S. Carey, B. S. Gaudi, C. B. Henderson, W. Zhu, Spitzer Team, M. D. Albrow, S. M. Cha, S. J. Chung, C. Han, K. H. Hwang, Y. K. Jung, D. J. Kim, H. W. Kim, S. L. Kim, Y. Lee, B. G. Park, R. W. Pogge, Y. H. Ryu, I. G. Shin, and KMTNet Group
2017. An Earth-mass Planet in a 1 au Orbit around an Ultracool Dwarf. *ApJL*, 840(1):L3.

Skidmore, W., TMT International Science Development Teams, and T. Science Advisory Committee
2015. Thirty Meter Telescope Detailed Science Case: 2015. *Research in Astronomy and Astrophysics*, 15(12):1945.

Steffen, J. H., D. C. Fabrycky, E. Agol, E. B. Ford, R. C. Morehead, W. D. Cochran, J. J. Lissauer, E. R. Adams, W. J. Borucki, S. Bryson, D. A. Caldwell, A. Dupree, J. M. Jenkins, P. Robertson, J. F. Rowe, S. Seader, S. Thompson, and J. D. Twicken
2013. Transit timing observations from Kepler - VII. Confirmation of 27 planets in 13 multiplanet systems via transit timing variations and orbital stability. *MNRAS*, 428(2):1077–1087.

Stewart, S. T. and Z. M. Leinhardt
2009. Velocity-Dependent Catastrophic Disruption Criteria for Planetesimals. *ApJ*, 691:L133–L137.

Stewart, S. T. and Z. M. Leinhardt
2012. Collisions between Gravity-dominated Bodies. II. The Diversity of Impact Outcomes during the End Stage of Planet Formation. *ApJ*, 751:32.

Struve, O.
1952. Proposal for a project of high-precision stellar radial velocity work. *The Observatory*, 72:199–200.

Swedenborg, E.
1734. *Principia*, Emanuelis Swedenborgii Opera philisophica et mineralia (English: Philosophical and Mineralogical Works). sumptibus Friderici Hekelii.

Tabachnik, S. A. and N. W. Evans
2000. Asteroids in the inner Solar system - I. Existence. *MNRAS*, 319:63–79.

Takeuchi, T., C. J. Clarke, and D. N. C. Lin
2005. The Differential Lifetimes of Protostellar Gas and Dust Disks. *ApJ*, 627(1):286–292.

Tanaka, H., T. Takeuchi, and W. R. Ward
2002. Three-Dimensional Interaction between a Planet and an Isothermal Gaseous Disk. I. Corotation and Lindblad Torques and Planet Migration. *ApJ*, 565(2):1257–1274.

Tanaka, H. and W. R. Ward
2004. Three-dimensional Interaction between a Planet and an Isothermal Gaseous Disk. II. Eccentricity Waves and Bending Waves. *ApJ*, 602(1):388–395.

Terquem, C. and J. C. B. Papaloizou
2007. Migration and the Formation of Systems of Hot Super-Earths and Neptunes. *ApJ*, 654(2):1110–1120.

Thompson, S. E., J. L. Coughlin, K. Hoffman, F. Mullally, J. L. Christiansen, C. J. Burke, S. Bryson, N. Batalha, M. R. Haas, J. Catanzarite, J. F. Rowe, G. Barentsen, D. A. Caldwell, B. D. Clarke, J. M. Jenkins, J. Li, D. W. Latham, J. J. Lissauer, S. Mathur, R. L. Morris, S. E. Seader, J. C. Smith, T. C. Klaus, J. D. Twicken, J. E. Van Cleve, B. Wohler, R. Akeson, D. R. Ciardi, W. D. Cochran, C. E. Henze, S. B. Howell, D. Huber, A. Prša, S. V. Ramírez, T. D. Morton, T. Barclay, J. R. Campbell, W. J. Chaplin, D. Charbonneau, J. Christensen-Dalsgaard, J. L. Dotson, L. Doyle, E. W. Dunham, A. K. Dupree, E. B. Ford, J. C. Geary, F. R. Girouard, H. Isaacson, H. Kjeldsen, E. V. Quintana, D. Ragozzine, M. Shabram, A. Shporer, V. Silva Aguirre, J. H. Steffen, M. Still, P. Tenenbaum, W. F. Welsh, A. Wolfgang, K. A. Zamudio, D. G. Koch, and W. J. Borucki
2018. Planetary Candidates Observed by Kepler. VIII. A Fully Automated Catalog with Measured Completeness and Reliability Based on Data Release 25. *The Astrophysical Journal Supplement Series*, 235:38.

Toomre, A.
1964. On the gravitational stability of a disk of stars. *ApJ*, 139:1217–1238.

Tremaine, S. and S. Dong
2012. The Statistics of Multi-planet Systems. *AJ*, 143:94.

Tsiaras, A., I. P. Waldmann, G. Tinetti, J. Tennyson, and S. N. Yurchenko
2019. Water vapour in the atmosphere of the habitable-zone eight-Earth-mass planet K2-18 b. *Nature Astronomy*, 3:1086–1091.

Udalski, A.
2003. The Optical Gravitational Lensing Experiment. Real Time Data Analysis Systems in the OGLE-III Survey. *Acta Astronomica*, 53:291–305.

Udry, S., M. Mayor, D. Naef, F. Pepe, D. Queloz, N. C. Santos, and M. Burnet
2002. The CORALIE survey for southern extra-solar planets. VIII. The very low-mass companions of <ASTROBJ>HD 141937</ASTROBJ>, <ASTROBJ>HD 162020</ASTROBJ>, <ASTROBJ>HD 168443</ASTROBJ> and <ASTROBJ>HD 202206</ASTROBJ>: Brown dwarfs or "superplanets"? *A&A*, 390:267–279.

Valencia, D., M. Ikoma, T. Guillot, and N. Nettelmann
2010. Composition and fate of short-period super-Earths. The case of CoRoT-7b. *A&A*, 516:A20.

Van de Kamp, P.
1969. Alternate dynamical analysis of Barnard's star. *AJ*, 74:757–759.

Van Eylen, V., C. Agentoft, M. S. Lundkvist, H. Kjeldsen, J. E. Owen, B. J. Fulton, E. Petigura, and I. Snellen
2018. An asteroseismic view of the radius valley: stripped cores, not born rocky. *MNRAS*, 479(4):4786–4795.

Van Eylen, V., S. Albrecht, X. Huang, M. G. MacDonald, R. I. Dawson, M. X. Cai, D. Foreman-Mackey, M. S. Lundkvist, V. Silva Aguirre, I. Snellen, and J. N. Winn
2019. The Orbital Eccentricity of Small Planet Systems. *AJ*, 157(2):61.

Visser, R. G. and C. W. Ormel
2016. On the growth of pebble-accreting planetesimals. *A&A*, 586:A66.

Vogt, S. S., S. L. Allen, B. C. Bigelow, L. Bresee, B. Brown, T. Cantrall, A. Conrad, M. Couture, C. Delaney, H. W. Epps, D. Hilyard, D. F. Hilyard, E. Horn, N. Jern, D. Kanto, M. J. Keane, R. I. Kibrick, J. W. Lewis, J. Osborne, G. H. Pardeilhan, T. Pfister, T. Ricketts, L. B. Robinson, R. J. Stover, D. Tucker, J. Ward, and M. Z. Wei
1994. HIRES: the high-resolution echelle spectrometer on the Keck 10-m Telescope. In *Instrumentation in Astronomy VIII*, D. L. Crawford and E. R. Craine, eds., volume 2198 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, P. 362.

Vogt, S. S., J. Burt, S. Meschiari, R. P. Butler, G. W. Henry, S. Wang, B. Holden, C. Gapp, R. Hanson, P. Arriagada, S. Keiser, J. Teske, and G. Laughlin
2015. Six Planets Orbiting HD 219134. *ApJ*, 814(1):12.

Vogt, S. S., R. P. Butler, J. Burt, M. Tuomi, G. Laughlin, B. Holden, J. K. Teske, S. A. Shectman, J. D. Crane, M. Díaz, I. B. Thompson, P. Arriagada, and S. Keiser
2017. A Six-planet System around the Star HD 34445. *AJ*, 154(5):181.

Ward, W. R.
1988. On disk-planet interactions and orbital eccentricities. *Icarus*, 73(2):330–348.

Weiss, L. M. and G. W. Marcy
2014. The Mass-Radius Relation for 65 Exoplanets Smaller than 4 Earth Radii. *ApJL*, 783:L6.

Weiss, L. M., G. W. Marcy, J. F. Rowe, A. W. Howard, H. Isaacson, J. J. Fortney, N. Miller, B.-O. Demory, D. A. Fischer, E. R. Adams, A. K. Dupree, S. B. Howell, R. Kolbl, J. A. Johnson, E. P. Horch, M. E. Everett, D. C. Fabrycky, and S. Seager
2013. The Mass of KOI-94d and a Relation for Planet Radius, Mass, and Incident Flux. *ApJ*, 768(1):14.

Wetherill, G. W. and G. R. Stewart
1989. Accumulation of a swarm of small planetesimals. *Icarus*, 77(2):330–357.

Wiegert, P. A. and M. J. Holman
1997. The Stability of Planets in the Alpha Centauri System. *AJ*, 113:1445–1450.

Wisdom, J. and M. Holman
1991. Symplectic maps for the N-body problem. *AJ*, 102:1528–1538.

Wolfgang, A., L. A. Rogers, and E. B. Ford
2016. Probabilistic Mass-Radius Relationship for Sub-Neptune-Sized Planets. *ApJ*, 825:19.

Wolszczan, A. and D. A. Frail
1992. A planetary system around the millisecond pulsar PSR1257 + 12. *Nature*, 355:145–147.

Wu, D.-H., R. C. Zhang, J.-L. Zhou, and J. H. Steffen
2018. Dynamical instability and its implications for planetary system architecture. *ArXiv e-prints*, P. arXiv:1809.08499.

Wu, D.-H., R. C. Zhang, J.-L. Zhou, and J. H. Steffen
2019. Dynamical instability and its implications for planetary system architecture. *MNRAS*, 484(2):1538–1548.

Wu, Y. and Y. Lithwick
2013. Density and Eccentricity of Kepler Planets. *ApJ*, 772:74.

Xie, J.-W.
2014. Transit Timing Variation of Near-resonance Planetary Pairs. II. Confirmation of 30 Planets in 15 Multiple-planet Systems. *ApJS*, 210(2):25.

Xie, J.-W., S. Dong, Z. Zhu, D. Huber, Z. Zheng, P. De Cat, J. Fu, H.-G. Liu, A. Luo, Y. Wu, H. Zhang, H. Zhang, J.-L. Zhou, Z. Cao, Y. Hou, Y. Wang, and Y. Zhang
2016. Exoplanet orbital eccentricities derived from LAMOST-Kepler analysis. *Proceedings of the National Academy of Science*, 113(41):11431–11435.

Yalinewich, A. and H. Schlichting
2019. Atmospheric mass-loss from high-velocity giant impacts. *MNRAS*, 486(2):2780–2789.

Youdin, A. N. and J. Goodman
2005. Streaming Instabilities in Protoplanetary Disks. *ApJ*, 620(1):459–469.

Zhou, J.-L., D. N. C. Lin, and Y.-S. Sun
2007. Post-oligarchic Evolution of Protoplanetary Embryos and the Stability of Planetary Systems. *ApJ*, 666(1):423–435.

Zhu, W., C. Petrovich, Y. Wu, S. Dong, and J. Xie
2018. About 30% of Sun-like Stars Have Kepler-like Planetary Systems: A Study of Their Intrinsic Architecture. *ApJ*, 860(2):101.

Zhu, W. and Y. Wu
2018. The Super Earth-Cold Jupiter Relations. *AJ*, 156(3):92.

Zink, J. K., J. L. Christiansen, and B. M. S. Hansen
2019. Accounting for incompleteness due to transit multiplicity in Kepler planet occurrence rates. *MNRAS*, 483(4):4479–4494.

List of Figures

List of Tables